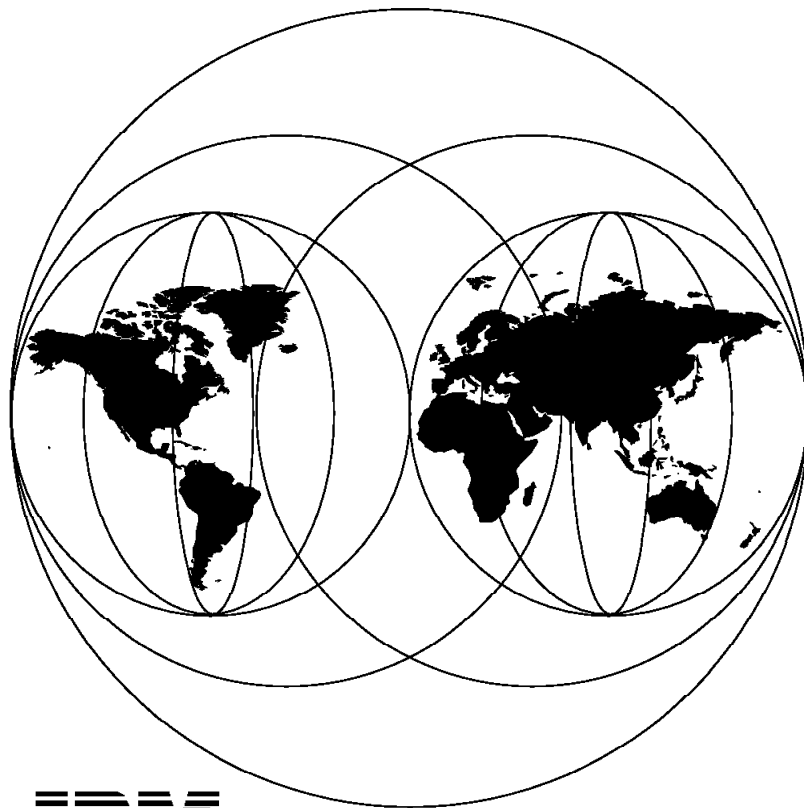


IBM RAMAC 3 Array Storage

December 1996



IBM

**International Technical Support Organization
San Jose Center**



International Technical Support Organization

SG24-4835-00

IBM RAMAC 3 Array Storage

December 1996

Take Note!

Before using this information and the product it supports, be sure to read the general information in Appendix B, "Special Notices" on page 169.

First Edition (December 1996)

This edition applies to the IBM 3990-6 and RAMAC 3 Array Storage.

Comments may be addressed to:

IBM Corporation, International Technical Support Organization
Dept. 471 Building 80-E2
650 Harry Road
San Jose, California 95120-6099

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

© **Copyright International Business Machines Corporation 1996. All rights reserved.**

Note to U.S. Government Users — Documentation related to restricted rights — Use, duplication or disclosure is subject to restrictions set forth in GSA ADP Schedule Contract with IBM Corp.

Contents

Preface	ix
How This Redbook Is Organized	ix
The Team That Wrote This Redbook	x
Comments Welcome	xi
 Chapter 1. Introduction	 1
Acknowledgments	2
Agenda	5
 Chapter 2. RAMAC 3 Array Storage Overview	 7
RAMAC 3 Array Storage Overview	7
RAMAC 3 Array Storage Overview	8
45.4 GB to 726.4 GB	8
256 MB to 8 GB of cache	8
32 MB to 256 MB of NVS	9
256 devices	9
4 to 32 ESCON channels	9
11.9 MB/s lower interface	9
General availability October 25, 1996	9
RAMAC 3 Array Storage Capacity	10
RAMAC 3 Array Storage - Solving the Puzzle	11
IBM 3990-6 and RAMAC 3 Storage Frame	12
IBM 3990-6 and IBM RAMAC 3 Storage Frame.. . . .	13
RAMAC 3 Frame attaches to a 3990-6	13
Requires LIC change	13
Supports 128 devices	13
Up to 363.2 GB	14
3390 and 3380 track format	14
Recommended minimum cache and NVS sizes	14
High-speed device adapter	14
RAMAC Array Family	15
 Chapter 3. IBM RAMAC 3 Array Drawer	 17
IBM RAMAC 3 Drawer	17
IBM RAMAC 3 Drawer...	18
RAMAC 3 drawer design	18
Enhanced drawer processing	18
High-Capacity Enabling	19
Performance	19
IBM RAMAC 3 Drawer	20
Common technology with RAMAC, RAMAC 2, and RAMAC 3	20
Proven high-availability RAID-5 design	20
Nonvolatile cache	20
22.7 GB per drawer	21
High performance SCSI 3.5 in. Ultrastar disk drives	21
RAMAC 3 Drawer...	22
Eight logical volumes per drawer	22
Four Ultrastar 2XP disks per drawer	22
High-speed parallel device adapter-device (PDAD)	23
Cache size of 64 MB	23
Improved data integrity	23

Hardware assist for fast response to initial volume access request	23
Improved sequential processing	23
RAMAC 3 Drawer Capacity	24
RAMAC 3 Drawer Logical Volume Emulation	25
RAMAC Drawer Logical Volume Mapping	26
RAMAC 2 Drawer Logical Volume Mapping	28
RAMAC 3 Drawer Logical Volume Mapping	30
RAMAC 3 Array Drawer Technology	31
RAMAC 3 Array Drawer Technology	33
RAMAC Drawer - System Components	35
Array management	35
Physical interface management	36
Lower interface data transfer management	36
Interprocessor communications	36
Cache management	36
Logical interface management	36
Background processing	37
Resource management	37
RAMAC 3 Drawer Hardware Assists	38
RAMAC 3 Drawer Hardware Assists...	39
RAMAC 3 Drawer Hardware Assists...	40
RAMAC Data Integrity and Availability	41
Disk scrubbing	41
RAMAC 3 Improved Data Integrity	42
SCSI header checking on every block	42
RAMAC 3 Improved Data Integrity...	43
Improved Sequential Write Processing	44
RAMAC Drawer Sequential Prestaging	46
Requested track and all tracks to end of cylinder...	46
Additional six tracks staged to "keep ahead"...	47
Drawer speed-matches overlapped data transfer	47
IBM RAMAC 3 Drawer - Summary	48
Highly parallel RAID-5 operations	48
Faster PDAD	48
Faster h/w-assisted prefetch	49
Sequential detect support	49
Advanced destage management	49
Logical volume mapping	49
High-performance Ultrastar 2XP disk drive	49
High-capacity Ultrastar 2XP disk drive	49
Eight logical volume emulation support	50
H/W-assisted SCSI block checking	50
CRC across CU-drawer interface	50
Third generation RAID-5	50
Fault-tolerant design	50
Chapter 4. IBM Ultrastar 2XP Disk Drive	51
IBM Ultrastar 2XP Disk Drive	52
9.22 GB	52
7200 rpm	52
No-ID sector format	53
Zoned bit recording	53
MR heads	53
Predictive Failure Analysis - S.M.A.R.T compliant	53
RAMAC 3 - Zoned Bit Recording	54

ID Sector Format	56
ID field needed to identify data	56
Presence of ID field reduces track density with MR head	57
Presence of ID field reduces track capacity	57
MR Head Geometry and ID Sector Format	58
MR Head Effect on ID Sector Format	60
No-ID Sector and ID Sector Format	61
Ultrastar 2XP No-ID Sector Format	62
ID information computed from disk drive solid state memory	62
Servo and data control systems perform high-speed address translation	62
Eliminates read ID before read and write operations	63
Advantages of Ultrastar 2XP No-ID Sector Format	64
Increased disk capacity	64
Improved performance	65
Improved defect management	65
Improved power management	65
Enhanced Reliability	65
RAMAC 3 Disk Drive Cache	66
1 MB disk drive cache	66
Cache segment size doubled	66
Drawer battery provides nonvolatility	66
Drawer and drive synergy	67
IBM Ultrastar 2XP Disk Drive - Performance	68
20 MB/s Fast-and-Wide SCSI	68
Media data rate 10.3-15.4 MB/s	68
Average seek (read) 8.5 ms	69
Track to track seek 0.5 ms	69
Latency 4.17 ms	69
IBM Ultrastar 2XP Disk Drive Reliability	70
PFA	70
Channel Calibration	70
Ultrastar 2XP Reliability...	71
Log recording	71
Disk sweep	71
IBM Ultrastar 2XP Disk Drive - Summary	72
High performance	72
High capacity	72
High reliability	73
Operating cost	73
Chapter 5. IBM RAMAC 3 Array Storage	75
IBM RAMAC 3 Array Storage	75
9390-001 Storage Control	76
Single Storage Control	76
Supports 128 devices	76
Cache	76
Nonvolatile storage	77
Channels	77
Basic mode default	77
9390-002 Storage Control	78
Dual storage control	78
Supports all extended functions	78
Supports 256 devices	79
Field upgrade from 9390-001 to 9390-002	79
Cache	79

Nonvolatile storage	79
Channels	79
Common power supplies	79
RAMAC 3 Storage Frame	80
New frame-9391-A30	80
9392-B33 drawer only	80
Minimum 2 drawers	80
Maximum 16 drawers	81
Attaches to 9390 or 3990-6	81
Minimum Configuration: 45.4 GB	82
Maximum Configuration: 726.4 GB	83
RAMAC 3 Addressing	84
9390 and 3990-6 can be set up to support:	84
RAMAC 3 Addressing: 64 Devices	85
RAMAC 3 Addressing: 128 Devices	86
 Chapter 6. 9390 and 3990-6 Storage Controls	 87
9390 and 3990-6 Storage Controls	87
High-Speed Device Adapter	88
HSDA cards	88
DDC cable	89
Effects of High-Speed Device Adapter	90
Faster destage from subsystem cache to drawer cache	90
Faster stage from drawer cache to subsystem cache	91
Faster PPRC initial volume synchronization	91
Faster XRC initial volume synchronization	91
Faster concurrent copy session management	91
New and faster storage frame maintenance through FPGA technology	92
No Changes to Storage Control hardware	92
Sequential Detect	93
Sequential detect is available on both the 3990-6 and 9390	93
Sequential detect is applicable to:	94
Sequential detect is invoked if more than three cylinders are read sequentially	94
Records are left in cache for normal LRU processing to allow reuse	94
VPD option to turn off: default is ON	94
Enhanced NVS Management	95
Current 3990-6 NVS management	95
Enhanced NVS management	96
Enhanced NVS Management...	97
Takes advantage of faster lower interface	97
Exploited by sequential detect	97
Uses nonvolatile drawer cache for second copy	98
Designed for long chain or large block size sequential data streams	98
Reduces NVS-full probability	98
Avoids asynchronous destage to drawer	98
Improved performance for mixed batch and online	98
Used for other writes if NVS becomes full	98
RAMAC 3 LIC Download and Activation	99
3990-6 LIC upgrade is disruptive	99
9390-001 and 9390-002	99
9390-2	99
RAMAC 3 Drawer Maintenance	100
RAMAC 3 Sparing	101
RAMAC 3 Drawer Maintenance - Summary	103

Dynamic Sparing preferred	103
Dynamic disk reconstruction	104
Chapter 7. Installation Planning	105
Installation and Migration	106
Software dependencies	106
IOCP definitions	107
Physical planning	107
VPD settings	107
ESCON connections	107
3990-6 support for IBM RAMAC 3 Array Storage	107
Data migration	107
9390-001 IOCP Definition Example	108
Two-CNTLUNIT macro definition	108
Single-CNTLUNIT macro definition	109
9390-001/ESCD IOCP Definition Example	110
ESCON Director Definition	110
9390/3990-6 Definition	111
9390-002 IOCP Definition Example	112
9390-002/ESCD IOCP Definition Example	113
Physical Specifications	114
Power and Floor Space Comparison	115
9390-001 Power Boundaries	116
9390-002 Power Boundaries	117
9390 Host Support	118
ES/9000, S/390, and 3090 processors with ESCON channels	118
ESCON support	118
9390-001 ESCON Channel Configuration	120
9390-002 ESCON Channel Configuration...	121
9390 Cache and NVS Feature Codes	122
3990-6 Support for RAMAC 3 Array	123
Customer investment protection	123
Faster DDC interface	123
New LIC download and activation required for 128 device support	123
Chapter 8. RAMAC 3 Array Storage and Remote Copy	125
9390 XRC Session Support	126
9390-002 at XRC Recovery Site	127
RAMAC 3 Array Storage and PPRC	129
9390-002 at PPRC Recovery Site	130
Chapter 9. Performance	131
Performance Tools	132
DCAT	132
CAA	132
DASD Magic/2	133
DSCAT	133
Performance with Sequential Detect	134
Sequential Performance Improvements	136
RAMAC 3 Recommended Configurations	137
Performance - Migration Consolidation	138
Chapter 10. Software	139
RAMAC 3 Array Storage Software Support	140
RAMAC 3 Software - APARs and PTFs	141

RAMAC 3 Software - APARs and PTFs...	142
RAMAC 3 Software - APARs and PTFs...	143
RAMAC 3 Software - APARs and PTFs...	144
Chapter 11. Operational Considerations	145
RAMAC 3 DEVSERV Format	146
Chapter 12. Positioning	147
RAMAC 3 Array Storage Positioning	148
RAMAC 3 Array Storage is designed for:	148
Positioning of New RAMAC Array Products	149
RAMAC 3 Array Storage	149
RAMAC Virtual Array	149
RAMAC Scalable Array	150
Positioning of RAMAC Family...	151
RAMAC 2 Array Subsystem	151
RAMAC Electronic Array	151
S/390 Internal Disk	151
IBM RAID Data Storage Offerings for S/390	152
Appendix A. RAMAC Advanced Destage Management	155
RAMAC Advanced Destage Management	155
RAMAC Advanced Destage Management...	157
RAMAC Advanced Destage Management...	158
RAMAC Advanced Destage Management...	159
RAMAC Advanced Destage Management...	160
RAMAC Advanced Destage Management...	161
RAMAC Advanced Destage Management...	163
RAMAC Advanced Destage Management...	164
Advanced Destage Management - Summary	166
Write penalty avoidance	166
Disk drive optimization	167
Storage control-RAMAC drawer synergy	167
Exploit drawer operating system	167
Appendix B. Special Notices	169
Appendix C. Related Publications	171
International Technical Support Organization Publications	171
Redbooks on CD-ROMs	171
Other Publications	171
How To Get ITSO Redbooks	173
How IBM Employees Can Get ITSO Redbooks	173
How Customers Can Get ITSO Redbooks	174
IBM Redbook Order Form	175
Index	177

Preface

This book describes the new RAMAC 3 Array Storage subsystem announced in September 1996. It provides detailed information about the advanced technology and architecture of the RAMAC 3 Array drawer and the Ultrastar 2XP disk drives. It provides guidance on installation planning and operations monitoring enhancements that assist customers in implementing RAMAC 3 Array Storage subsystems.

The book is written for storage administrators, systems programmers, site operations staff, and other technical support staff and managers interested in storage subsystem configuration design, capacity planning, performance, and installation and migration.

How This Redbook Is Organized

This redbook contains 210 pages and the following chapters:

- IBM RAMAC 3 Array Storage Overview

This chapter introduces the new RAMAC 3 Array Storage and provides an overview of its capacity, channel connectivity, and cache and NVS configurations.

- IBM RAMAC 3 Array Drawer

This chapter revisits the technology and design of the RAMAC and RAMAC 2 drawers and highlights the changes implemented in the RAMAC 3 drawer.

- IBM Ultrastar 2XP Disk Drive

This chapter describes the technology and architecture of the new IBM Ultrastar 2XP disk drive. The Ultrastar 2XP's effect on capacity, performance, and reliability is discussed.

- IBM RAMAC Array Storage

This chapter describes details of the RAMAC 3 Array Storage, including the new RAMAC 3 Storage Frame and its attachment to the new 9390-001 and 9390-002 storage controls and the 3990-6 storage control. Device addressability is also covered.

- 9390 and 3990-6 Storage Controls

The details of how the 9390 and 3990-6 storage controls support the new RAMAC 3 Array drawer and RAMAC 3 Storage Frame are presented in this chapter.

- Installation Planning

This chapter describes the key tasks involved in planning for installation of the new RAMAC 3 Array Storage. Tasks covered include software requirements, IOCP definitions, and physical and environmental specifications.

- RAMAC 3 Array Storage and Remote Copy

The RAMAC 3 Array Storage can participate in both XRC and PPRC remote copy configurations. This chapter describes details of how the attributes of the RAMAC 3 Array Storage affect XRC and PPRC configuration design.

- Performance

This chapter contains details of performance comparisons between 3990-6 and RAMAC 2 and RAMAC 3 Array Storage for specified workloads. The performance of RAMAC 3 Array Storage using the new sequential detect algorithms and the effect of migration and consolidation of capacity and workload are also described.

- Software

This chapter highlights the software support that enables implementation and exploitation of the RAMAC 3 Array Storage.

- Operational Considerations

The details that must be communicated to customer operations staff when implementing and monitoring RAMAC 3 Array Storage subsystems are described in this chapter.

- Positioning

RAMAC 3 Array Storage is joined by three other new members of the RAMAC Array family—the RAMAC Virtual Array, the RAMAC Scalable Array, and the RAMAC Electronic Array. This chapter describes the functions of the new members, and explains how one might be favored over another depending on installation requirements.

- Appendix A, “RAMAC Advanced Destage Management”

This appendix contains a step-by-step description of how the RAMAC 3 Array drawer manages the parallel data transfers and utilization of drawer components as part of its data destage activities.

The Team That Wrote This Redbook

This redbook was produced by a team of specialists from around the world working at the International Technical Support Organization San Jose Center.

Dave McAuley is the project leader for large systems disk storage products at the International Technical Support Organization, San Jose Center. He has written six ITSO redbooks and several international conference papers, among them a paper on image processing and pattern recognition in IBM's *Journal of Research and Development*. Dave teaches IBM classes worldwide on all areas of high-end DASD products. Before joining the ITSO in 1994, he worked in the product support center in Scotland as a large systems storage specialist. Dave has worked at IBM for 18 years, with international career assignments in manufacturing, design and development, and marketing. His background in computer science spans a varied portfolio including supercomputing application development, process control and image analysis systems design and microcode development, automation and robotics systems project management, and large systems marketing and technical support.

Phil Norman is an Information Technology consultant in the Enterprise Systems technical support group in the UK. He has 10 years of experience in the storage hardware field. His areas of expertise include 3990, 3390, and RAMAC hardware, and remote copy. Phil specializes in supporting performance studies, field guide preparation, and workshops.

Hans Esch is an information technology marketing specialist in Duesseldorf, Germany. He has more than 20 years of experience in S/370 and S/390 systems

and large systems storage technical support and marketing. His other areas of expertise include VM/ESA and VSE/ESA networking and connectivity. Hans joined IBM in 1969.

Thanks to the following people for their invaluable contributions to this project:

Dave Giese, Storage Systems Division, San Jose
Steve Hetzler, Almaden Research Center, San Jose
Joe Hyde, Storage Systems Division, Tucson
Scott Karstetter, Storage Systems Division, San Jose
Ted Lee, Storage Systems Division, San Jose
David Plomgren, Storage Systems Division, San Jose
Greg Sims, Storage Systems Division, San Jose
Jeff Steffan, Storage Systems Division, San Jose
Karl Stoeckli, Credit Suisse, Zurich, Switzerland
Chris Xydes, Storage Systems Division, San Jose

Thanks are also due to the many other collaborators and reviewers who contributed to the production of this document. In particular, we would like to thank the following people:

Maggie Cutler, ITSO Technical Editor
Jacques Glorieux, IBM Belgium
Dave Hollar, S/390 Division
Chuck Lanzi, Storage Systems Division, San Jose
John Power, Storage Systems Division, San Jose
Charlie Shapley, S/390 Division
Frank Vitro, Storage Systems Division, San Jose

Comments Welcome

We want our redbooks to be as helpful as possible. Should you have any comments about this or other redbooks, please send us a note at the following address:

redbook@vnet.ibm.com

Your comments are important to us!

Chapter 1. Introduction

Introduction



IBM RAMAC 3 Array Storage

Dave McAuley
International Technical Support Organization
San Jose

October 1996



© IBM Corporation 1996

Acknowledgments



- Project leader
 - Dave McAuley - ITSO-San Jose Center
- Residents
 - Hans Esch - IBM Germany
 - Phil Norman - IBM UK



© IBM Corporation 1996

Acknowledgments

This book is the result of a residency project run at the International Technical Support Organization-San Jose Center in San Jose, California, during July and August 1996. The project was designed and managed by:

Dave McAuley, International Technical Support Organization-San Jose Center

The book was written by:

- Dave McAuley, IBM ITSO-San Jose Center
- Hans Esch, IBM Germany
- Phil Norman, IBM UK

We are grateful for the contributions of many people, but particularly acknowledge:

- Dave Giese, Storage Systems Division, San Jose
- Steve Hetzler, Almaden Research Center, San Jose
- Joe Hyde, Storage Systems Division, Tucson
- Scott Karstetter, Storage Systems Division, San Jose
- Ted Lee, Storage Systems Division, San Jose

- David Plomgren, Storage Systems Division, San Jose
- Greg Sims, Storage Systems Division, San Jose
- Jeff Steffan, Storage Systems Division, San Jose
- Karl Stoeckli, Credit Suisse, Zurich, Switzerland
- Chris Xydes, Storage Systems Division, San Jose

Audience



- This presentation is written for:
 - Customer personnel
 - IBM technical and marketing specialists
 - IBM client representatives



© IBM Corporation 1996

Trademarks



The following products and features referenced in this presentation are trademarks of the IBM Corporation in the United States and/or other countries:

- | | |
|--|-------------------------------|
| ■ ES/9000 | ■ MVS/ESA |
| ■ ESCON | ■ MVS/XA |
| ■ IBM | ■ VM/ESA |
| ■ RAMAC | ■ VM/XA SP |
| ■ System/390 | ■ VSE/ESA |
| ■ Hardware Configuration
Definition | ■ DB2 |
| ■ MVS/DFP | ■ IMS/ESA |
| ■ DFSMS/MVS | ■ CICS/ESA |
| ■ No-ID Sector Format | ■ NetView |
| | ■ Predictive Failure Analysis |



© IBM Corporation 1996

Agenda



- Overview of RAMAC 3 Array Storage
- IBM RAMAC 3 drawer
- IBM Ultrastar 2XP disk drive
- IBM RAMAC 3 Storage Control and RAMAC 3 Storage Frame
- IBM RAMAC 3 Storage Control and 3990 Model 6
- Operational Considerations
- Performance
- Installation and Migration
- Software Support and Tools
- Positioning
- Summary

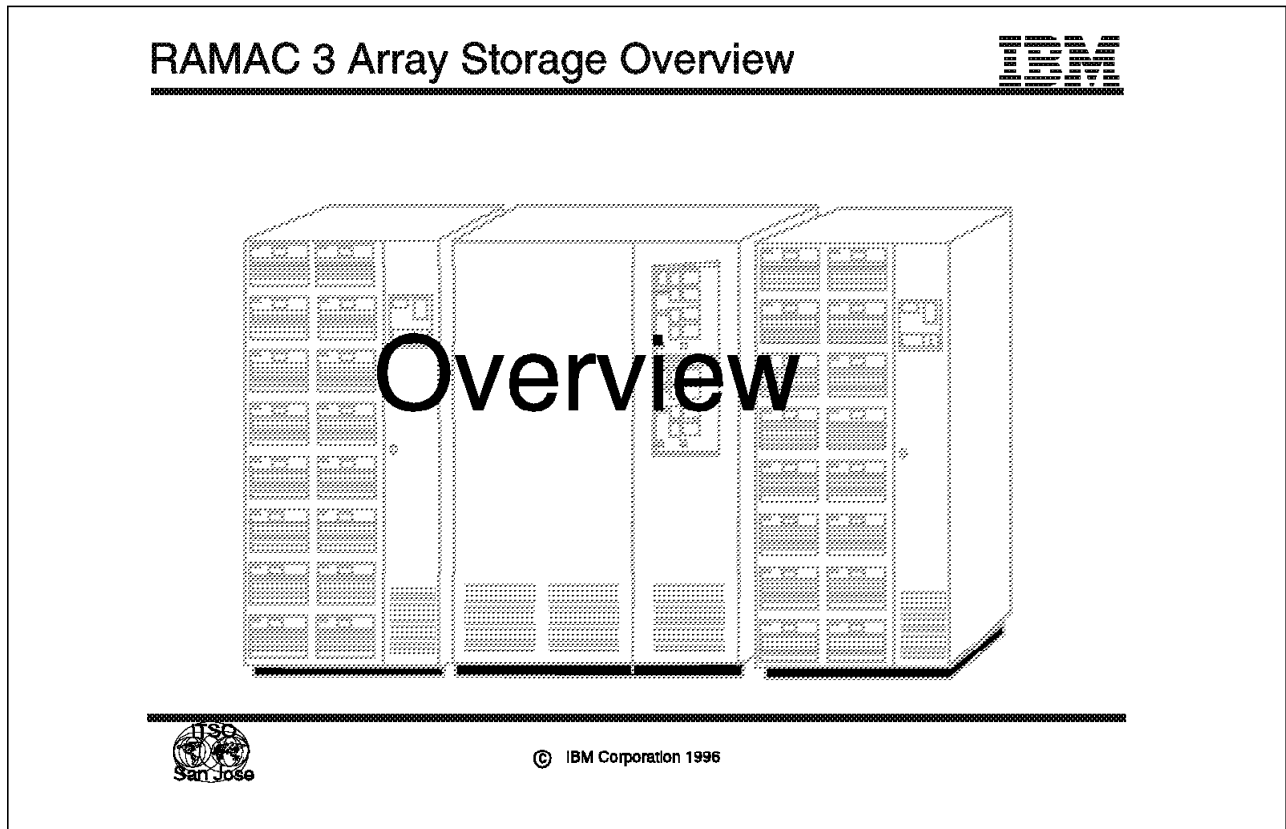


© IBM Corporation 1996

Agenda

This foil lists the topics covered in this presentation guide. We provide a brief overview of the new IBM RAMAC 3 Array Storage subsystem, the attachment of the RAMAC 3 Storage Frame to an existing 3990-6, and the functions that are supported. Then we describe each of the components of the RAMAC 3 Array Storage, from the drawer and new disk drive, to the enhanced device adapters and the new RAMAC 3 Storage Control. We review configuration options; operational, performance, installation, and migration considerations; and software support and tools. We position the new RAMAC 3 Array Storage in the enlarged RAMAC Array family, now consisting of RAMAC, RAMAC 2, RAMAC 3, the RAMAC Virtual Array, the RAMAC Scalable Array, and the RAMAC Electronic Array plus the new S/390 Internal Disk. We conclude with a summary of the new products and enhancements.

Chapter 2. RAMAC 3 Array Storage Overview

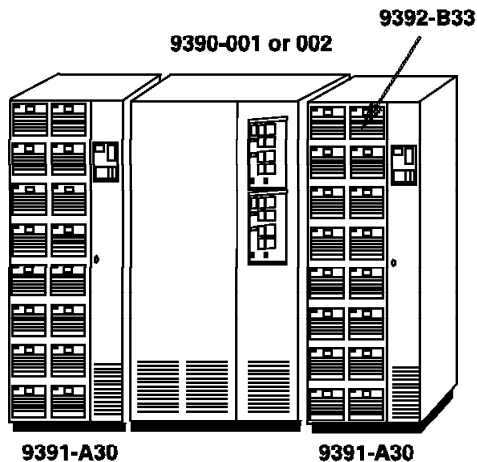


RAMAC 3 Array Storage Overview

This section gives a brief overview of the IBM RAMAC 3 Array Storage and its capacity and configuration options. The schematic shows the RAMAC 3 Storage Control and two RAMAC 3 Storage Frames. The storage control is the same height as a RAMAC frame and contains two storage controls in one cabinet. We also look at the attachment of the RAMAC 3 Storage Frame to an existing 3990-6 and position the new RAMAC 3 Array Storage in the RAMAC Array family of products.

RAMAC 3 Array Storage Overview...

- 45.4 GB to 726.4 GB
- 256 MB to 8 GB of cache
- 32 MB to 256 MB of NVS
- 256 devices
- 4 to 32 ESCON channels
- 11.9 MB/s lower interface
- General availability
October 25, 1996



© IBM Corporation 1996

RAMAC 3 Array Storage Overview ...

Here we summarize the capacity, connections, performance, and availability of the RAMAC 3 Array Storage.

45.4 GB to 726.4 GB

The IBM RAMAC 3 Array Storage subsystem consists of an IBM RAMAC Storage Control and one or two IBM RAMAC 3 Storage Frames each containing up to 363.2 GB of storage. A fully configured IBM RAMAC 3 Storage Frame contains 16 RAMAC 3 drawers, each with 22.7 GB of storage capacity, giving a total of 363.2 GB. Therefore, two RAMAC 3 Storage Frames each containing 16 drawers provide 726.4 GB. The minimum configuration is a RAMAC 3 Storage Frame containing two RAMAC 3 drawers, providing a total capacity of 45.4 GB.

There are two models of the RAMAC Storage Control: the IBM RAMAC Storage Control Model 001 (9390-001), which contains one storage control, and the IBM RAMAC Storage Control Model 002 (9390-002), which contains two storage controls. Each storage control connects to one of the Storage Frames.

256 MB to 8 GB of cache

The minimum cache size for a storage control is 256 MB. This size ensures that the storage subsystem delivers good performance for lower-capacity configurations. Maximum cache is 8 GB (4 GB per storage control in a 9390-002).

32 MB to 256 MB of NVS

The nonvolatile storage sizes vary from a minimum of 32 MB (on a 9390-001 single storage control) to 256 MB (128 MB per storage control on the 9390-002).

256 devices

Each storage control on a 9390-002 supports up to 128 devices for a total of 256. Each of the two RAMAC 3 frames contains from 2 to 16 drawers, each emulating eight logical 3390-3 volumes.

4 to 32 ESCON channels

Channel attachment is ESCON only; parallel channel support is not provided. The minimum channel configuration is 4 ESCON channels on a 9390-001 and the maximum is 32 on a 9390-002 (16 on each storage control).

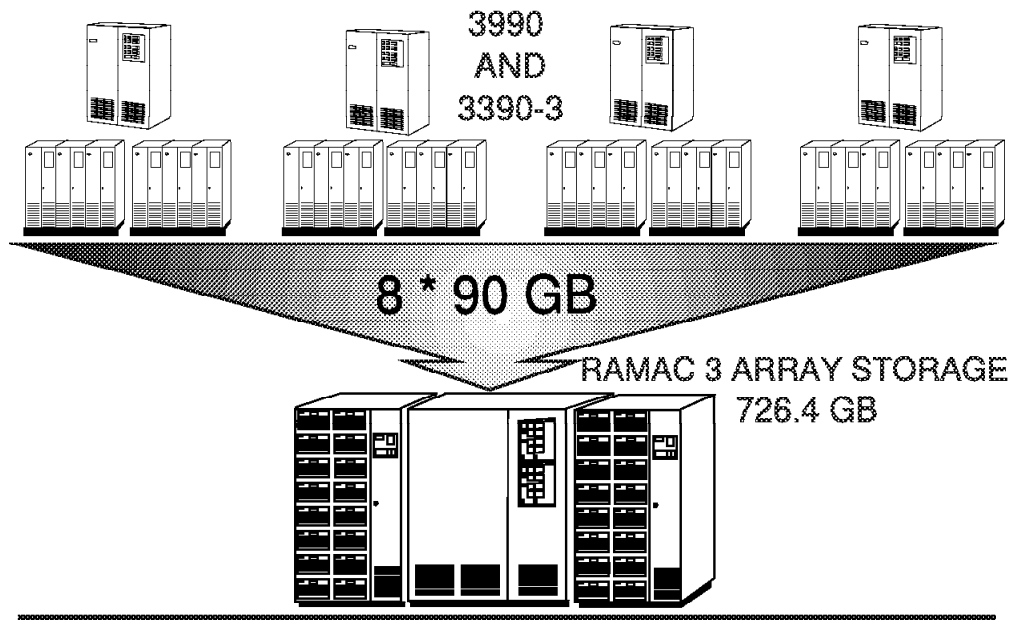
11.9 MB/s lower interface

The lower interface between the RAMAC 3 Storage Control and the RAMAC 3 Storage Frame has been enhanced to support a substantially higher data rate. The increase to 11.9 MB/s enables the storage control to support the larger number of logical volumes. In addition, the faster data rate benefits sequential operations and improves batch job throughput, particularly important when extended online hours squeeze the batch window.

General availability October 25, 1996

General availability is October 25, 1996.

RAMAC 3 Array Storage Capacity



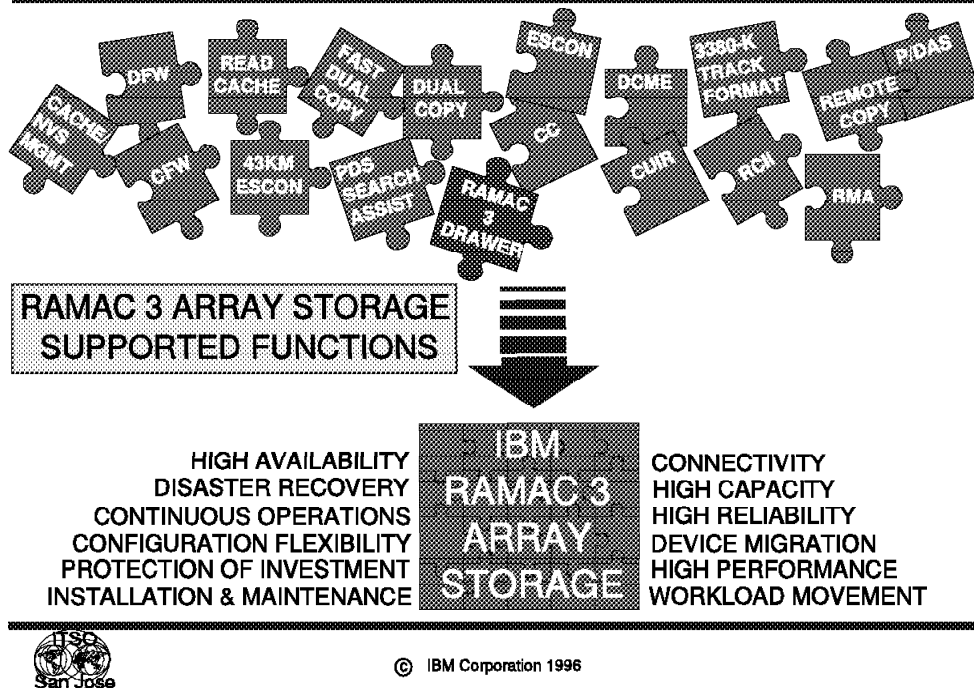
© IBM Corporation 1996

RAMAC 3 Array Storage Capacity

The RAMAC 3 Array Storage provides a dramatic reduction in floor space per gigabyte when compared with previous DASD subsystems. For example, a fully configured RAMAC 3 Array Storage subsystem has the equivalent capacity of four 3990s, each with two strings of 3390-3.

If we look at the capacity in terms of the 3380 and 3880, we have a physical reduction from about 192 to 3 boxes.

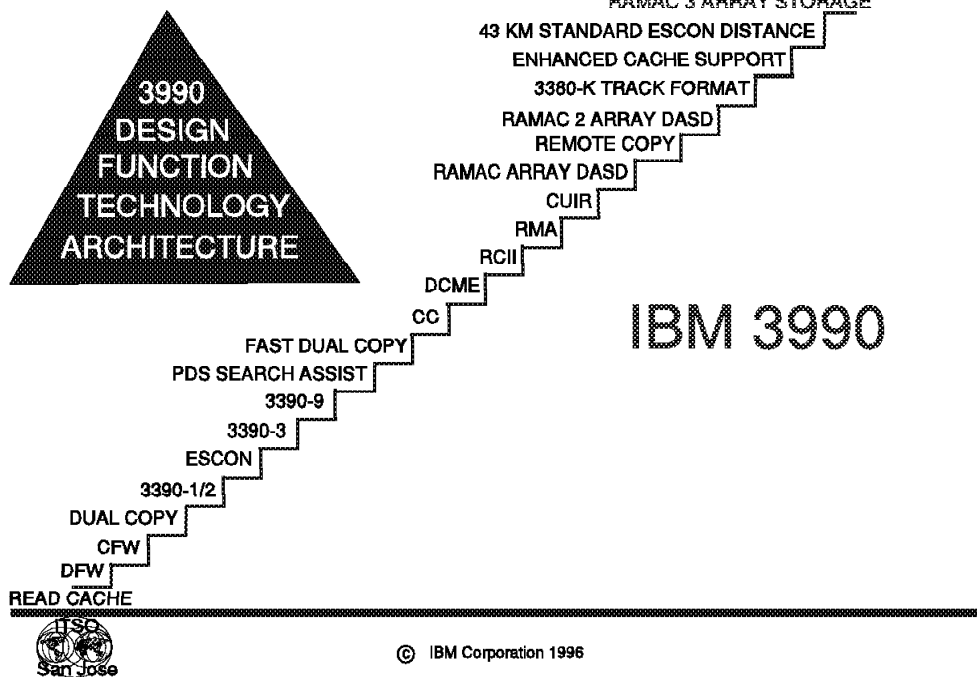
RAMAC 3 Array Storage-Solving the Puzzle



RAMAC 3 Array Storage - Solving the Puzzle

The RAMAC 3 Array Storage subsystem provides all of the extended functions and I/O support capability that have made the 3990-6 an outstanding storage control. The extended functions of the RAMAC 3 Array Storage address the key requirements of today's high-end DASD subsystem and are not simply a group of disparate features developed in isolation—instead, these functions "fit together" to deliver a solutions-oriented base for data storage and retrieval.

IBM 3990-6 and RAMAC 3 Storage Frame



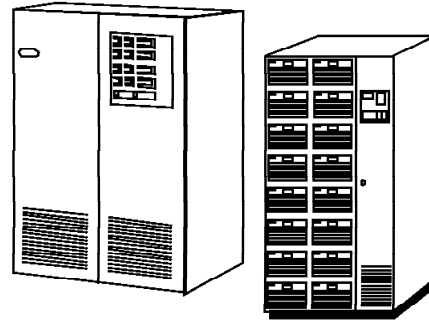
IBM 3990-6 and RAMAC 3 Storage Frame

The 3990-6 has been at the center of IBM's storage strategy for almost a decade, delivering high performance and high availability with an ongoing evolution of functions. With the announcement of 3990-6 support for the IBM RAMAC 3 Storage Frame, IBM is protecting your investment in installed controllers. All of the extended functions supported on the 3990-6 are enabled for RAMAC 3 Array DASD configurations including the unique support for both parallel and ESCON host channel attachments.

3990-6 and RAMAC 3 Storage Frame



- RAMAC 3 Storage Frame attaches to 3990-6
- Requires LIC change
- Supports 128 devices
- Up to 363.2 GB
- 3390 and 3380 track format
- Recommended minimum cache and NVS sizes
- High-speed device adapter



© IBM Corporation 1996

IBM 3990-6 and IBM RAMAC 3 Storage Frame..

RAMAC 3 Frame attaches to a 3990-6

All 3990-6 Storage Controls can support the new RAMAC 3 Storage Frame, although new director-to-device connection (DDC) cables are required. Some early models of 3990-6 require a larger capacity storage control hard drive.

Requires LIC change

A new level of Licensed Internal Code is required to provide the support for the new interface speeds and additional addresses when the 3990-6 attaches to the RAMAC 3 Storage Frame.

See Chapter 10, "Software" on page 139 for details on software support for the RAMAC 3 Array Storage.

Supports 128 devices

The additional capacity is provided by increasing the number of addresses from 64 to 128 in the 3990-6.

Up to 363.2 GB

With the increase in the number of volumes, the 3990 Model 6 can now support up to 363.2 GB in 3390-3 format.

3390 and 3380 track format

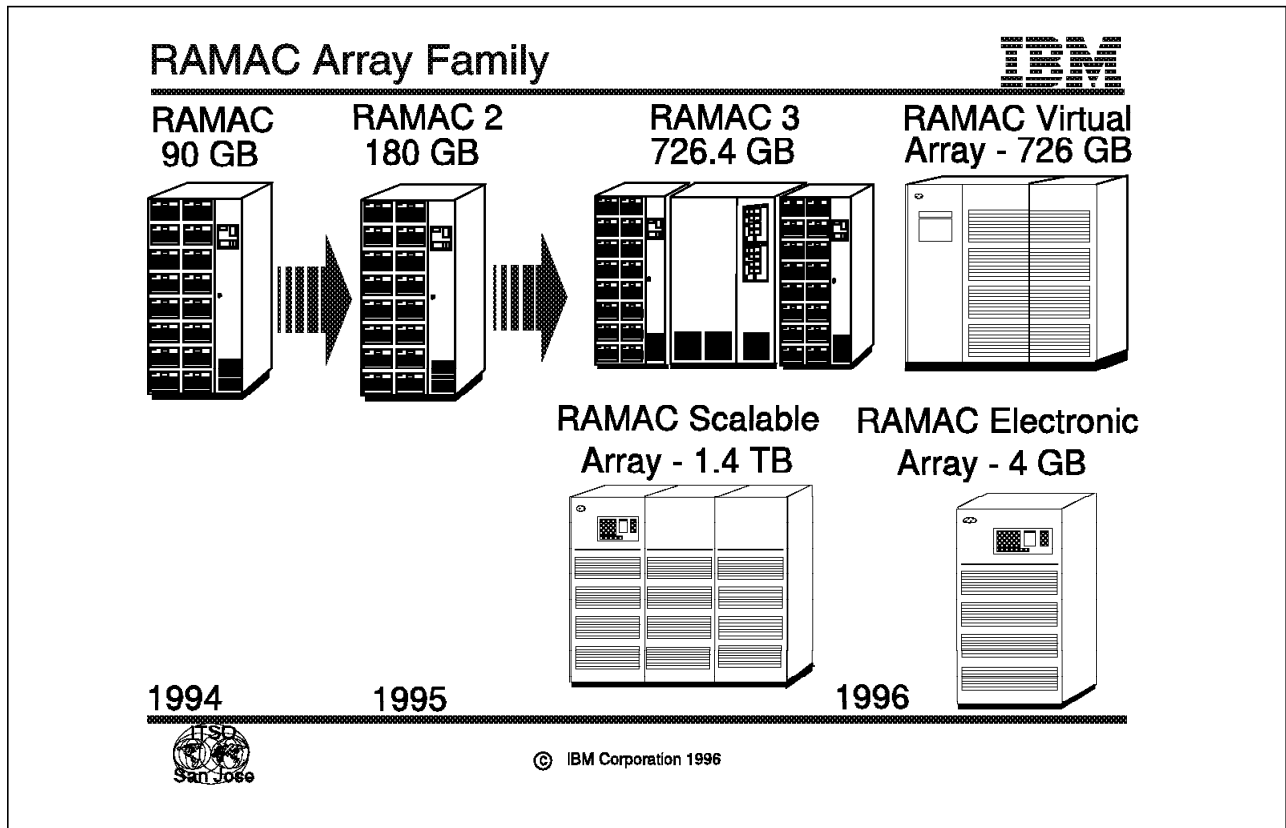
The 3990-6 supports both 3390 and 3380 track format. If a full RAMAC 3 Storage Frame is formatted in 3380 track format, the capacity is reduced by 16%, from 363.2 GB per rack to 304 GB.

Recommended minimum cache and NVS sizes

There are recommended minimum cache and NVS sizes for RAMAC 3 Storage Frame configurations. The minimum sizes for the 9390 with RAMAC 3 is 256 MB of cache and 32 MB of NVS; a similar recommendation applies to the 3990-6 when attaching a RAMAC 3 Storage Frame. A more typical size for high performance would be 1 GB of cache and 64 MB of NVS. We recommend that you review the cache size of an existing 3990 Model 6 with your IBM storage specialist if you plan to attach the new RAMAC 3 Storage Frame.

High-speed device adapter

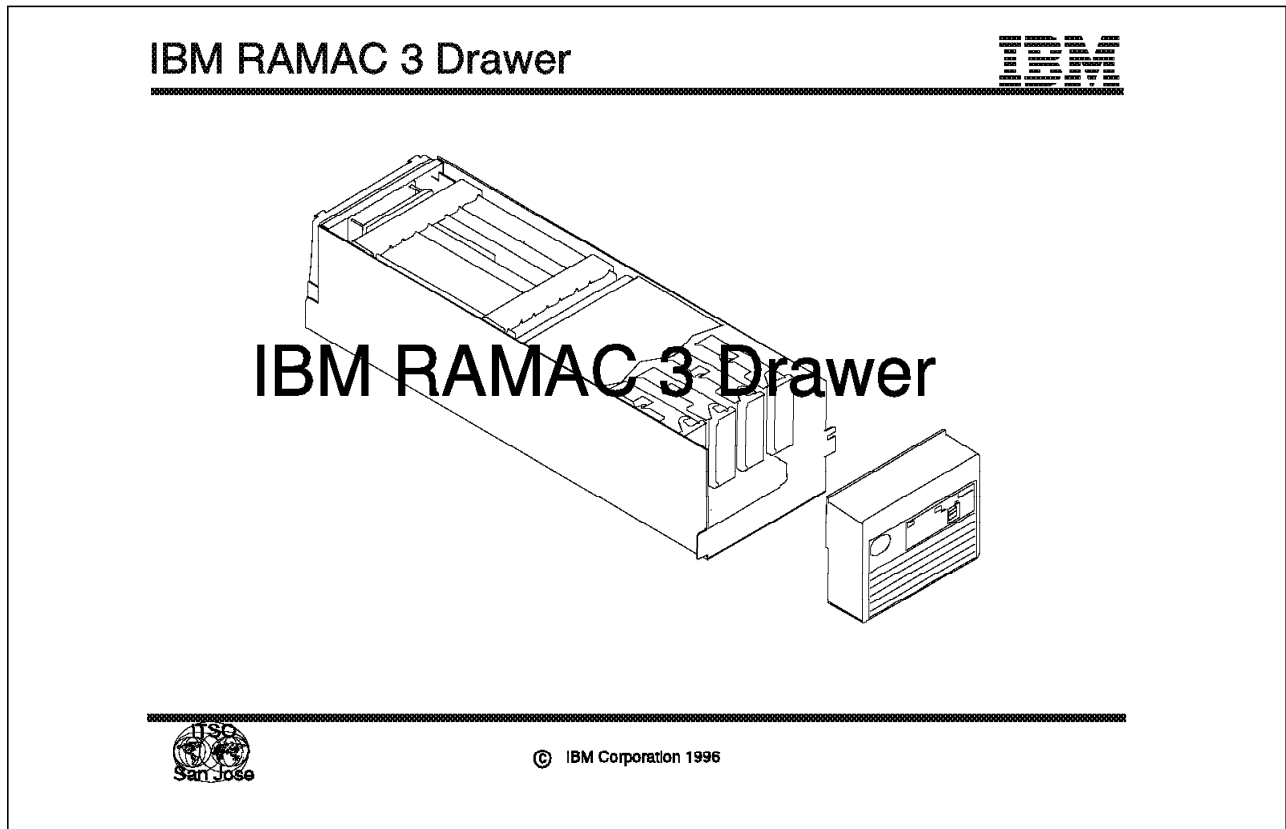
Until now the 3990-6 to RAMAC drawer lower interface has operated at the RAMAC and RAMAC 2 parallel device adapter (PDAD) speed. RAMAC 3 has a faster PDAD, triple the speed of the RAMAC and RAMAC 2 adapter at 11.9 MB/s. The faster interface also improves the sequential performance of the 3990-6 with RAMAC 3 drawers. This enhancement is available only with RAMAC 3 drawers and either a 9390 or a 3990-6.



RAMAC Array Family

The RAMAC Array Family has been significantly enhanced by the addition in 1996 of the RAMAC Virtual Array and the RAMAC Scalable Array (with the RAMAC Electronic Array as a specialized version of the Scalable Array) to the RAMAC 2 Array DASD and the RAMAC 2 Array Subsystem. Now, adding to the customer's choice of S/390 DASD, IBM introduces the new RAMAC 3 Array Storage. Each of the storage products provides high availability together with product-unique functions and characteristics.

Chapter 3. IBM RAMAC 3 Array Drawer



IBM RAMAC 3 Drawer

Apart from the larger capacity disks, there are many enhancements to the RAMAC 3 Array drawer. This section revisits the RAMAC drawer technology and design, and describes the changes made to the RAMAC 3 drawer.

IBM RAMAC 3 Drawer...



- RAMAC 3 drawer design
 - Hardware components
 - Software components
 - Recovery
- Enhanced drawer processing
 - Advanced destage management
 - Hardware-assisted functions
- High-capacity enabling
 - Volume emulation
- Performance
 - Delivered by:
 - *Hardware*
 - *Software*
 - *Architecture*
 - *Highly parallel design*
 - *Storage control synergy*
 - *Disk drive synergy*



© IBM Corporation 1996

IBM RAMAC 3 Drawer...

The IBM RAMAC Array Family of products implements a hierarchical, distributed design that results in highly parallel storage subsystem processing. The 3990 and 9390 storage controls attach to the RAMAC 3 Storage Frame, which in turn houses the RAMAC 3 drawer. The RAMAC 3 drawer contains the Ultrastar 2XP disk drives. Each of these components performs vital storage subsystem tasks in harmony, but also in parallel. This section focuses on the changes to the RAMAC 3 drawer.

RAMAC 3 drawer design

The RAMAC 3 drawer transforms standard RAID-5 architecture into advanced RAID-5 operations by exploiting a sophisticated “web” of hardware and microcode components that have matured and stabilized through three generations of the RAMAC family. These components and how they interact are described in this section.

Enhanced drawer processing

The RAMAC 3 drawer neutralizes many of the adverse performance effects associated with RAID-5 processing. From a performance perspective, write penalty avoidance is the major goal of RAID-5 (and other RAID implementations) storage subsystems and must be achieved without compromising data integrity and data availability. The RAMAC Array Family uses a combination of hardware and microcode to provide unique techniques such as advanced destage management and hardware-assisted data prefetch functions. These techniques are described in this section.

High-Capacity Enabling

The RAMAC Array Family of products has consistently delivered higher-capacity subsystems year after year. This answers the relentless demand created by the global explosion of information systems and networks. Throughout this period of growth the RAMAC Array Family has adhered to a technique that enables cost-effective, risk-free increases in the amount of data that can be accessed by an application. Logical volume emulation provides a secure and easy way to capitalize on the advances made in disk drive technology and architecture. RAMAC 3 Array Storage continues this technique by providing eight logical volumes in each RAMAC 3 drawer.

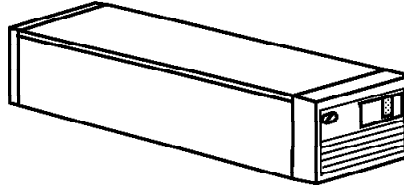
Performance

A single RAMAC 3 Array Storage configuration quadruples the storage capacity of a single 3990-3 storage control while maintaining the highest levels of performance for advanced RAID-5 operations. Performance improvements implemented in the RAMAC 3 Storage Frame and RAMAC 3 drawer are complemented by enhancements in drawer microcode and disk drive technology and architecture. The relationship between the RAMAC 3 drawer and the 9390 and 3990-6 storage controls also results in more efficient data transfer setup and parity processing.

IBM RAMAC 3 Drawer ...



- Common technology with RAMAC, RAMAC 2, and RAMAC 3
- Proven high-availability RAID-5 design
- Nonvolatile cache
- 22.7 GB per drawer
- High-performance SCSI 3.5 in. Ultrastar disk drives



© IBM Corporation 1996

IBM RAMAC 3 Drawer

The RAMAC 3 drawer enables more cost-effective capacity than the RAMAC 2 drawer through its use of the Ultrastar 2XP disk drive and improved drawer components.

Common technology with RAMAC, RAMAC 2, and RAMAC 3

The technology base and drawer design is common among RAMAC family members, thus enabling low-risk hardware and microcode enhancements on the RAMAC 3 drawer in addition to exploitation of the new Ultrastar 2XP disk drive.

Proven high-availability RAID-5 design

The enhanced RAID-5 design of the RAMAC 3 drawer builds on the proven high-availability RAID-5 design of the previous RAMAC drawers. RAMAC has been certified as compliant with the RAID Advisory Board standards for RAID-5 and is entitled to use the RAB logo.

Nonvolatile cache

In the multilevel cache design of RAMAC, where the RAID-5 handling is devolved to the drawer, the nonvolatile cache is used to hold data while parity is calculated. It is also used to prestage data from the disks to the drawer cache ready to be transferred to the storage control.

22.7 GB per drawer

The capacity of a RAMAC 3 drawer has been increased to 22.7 GB. This is four times the capacity of RAMAC and twice that of RAMAC 2. The additional capacity is derived from the use of the IBM Ultrastar 2XP disk drive that provides 9.22 GB of storage capacity. The Ultrastar disk drives are arranged in a RAID-5 array formation in the RAMAC 3 Array drawer.

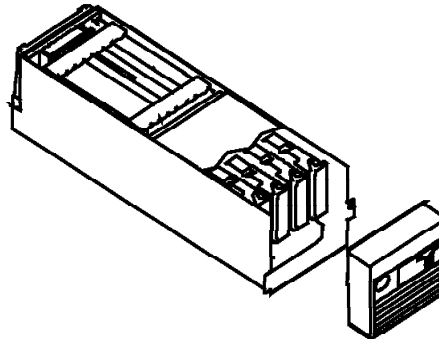
High performance SCSI 3.5 in. Ultrastar disk drives

The new disk drive in the RAMAC 3 Array drawer is an IBM Ultrastar 2XP. This disk drive, also used in the S/390 Internal Disk and in the enhanced RAMAC Virtual Array, demonstrates IBM's position as a leader in the disk technology industry.

IBM RAMAC 3 Drawer ...



- Eight logical volumes per drawer
- Four Ultrastar 2XP disks per drawer
- High-speed parallel device adapter-device (PDAD)
- Cache size of 64 MB
- Improved data integrity
- Hardware assist for fast response to initial volume access request
- Improved sequential processing



© IBM Corporation 1996

RAMAC 3 Drawer...

The RAMAC drawer consists of many components. Here we look in detail at the changes introduced in RAMAC 3 and in the enhancements that benefit RAMAC, RAMAC 2, and RAMAC 3.

Eight logical volumes per drawer

The double capacity of RAMAC 3 over RAMAC 2 is achieved by mapping eight 3390-3 volumes in each drawer. As in other RAMAC drawer designs, the tracks of each of the emulated disks are striped across three of the four disks, the fourth disk holding the parity track. As in any RAID-5 design, the parity rotates around the disks to spread the parity I/O activity across all disks.

Four Ultrastar 2XP disks per drawer

The new Ultrastar disks are located in the RAMAC 3 drawer in the same way as in the existing RAMAC drawers. A disk can be removed and replaced should it fail. The RAMAC drawer rebuilds the data and parity on the replaced disk automatically.

High-speed parallel device adapter-device (PDAD)

To support the larger capacity in each drawer and to enhance data transfer, staging, and destaging performance, the interface between the storage control and the RAMAC 3 Storage Frame and between the Storage Frame and the RAMAC 3 drawer has been redesigned. As part of the new design the PDAD in the RAMAC 3 drawer has been enhanced to handle a tripling of the data rate. Foil "High-Speed Device Adapter" on page 88 gives more details.

Cache size of 64 MB

The cache size of a RAMAC 3 drawer remains at 64 MB, the same as that of RAMAC and RAMAC 2. It may seem that a larger size would be necessary, particularly as we are still using the same size as RAMAC but with four times the disk capacity per drawer. Both the enhanced data rate on the upper drawer interface and the faster disk speed on the lower interface reduce the time data is held in the drawer cache and thereby the need for a larger cache. When the ratio between the 3990 cache and drawer cache was low, for example, when the 3990 had 64 MB of cache and the drawer 64 MB of cache, there were performance benefits from drawer cache hits. As the storage control caches have increased, drawer cache hit benefits have been reduced. With low hit rates in a drawer, the drawer cache is mainly used for RAID-5 parity handling and staging and destaging tracks. Analysis has shown that a larger storage control cache and NVS provide better overall performance than increasing the size of drawer cache and are more cost effective than increasing the cache size in every drawer. For recommendations on storage control cache and NVS sizes, see foil "RAMAC 3 Recommended Configurations" on page 137.

Improved data integrity

Since the introduction of RAMAC, IBM has consistently improved its data integrity and availability characteristics. Today RAMAC is the industry leading S/390 storage product in terms of reliability, availability, and data integrity. In each iteration of the RAMAC design, enhancements are made on the basis of practical experience with the products and their behavior in customer environments. In RAMAC 3 further enhancements have been made, as discussed in more detail on foil "RAMAC Data Integrity and Availability" on page 41.

Hardware assist for fast response to initial volume access request

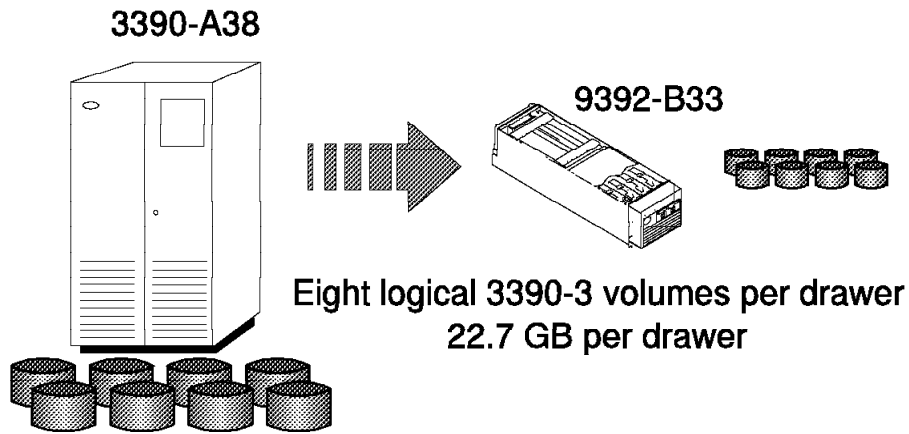
In the RAMAC drawer, many functions are implemented in the microcode rather than the hardware for both ease of change and cost. In some cases, however, a hardware implementation of a function can be justified on the significant performance improvement that it delivers. A number of functions have been implemented in the hardware in the RAMAC 3 drawer that complement the higher data rate from the storage control.

Improved sequential processing

A large part of your data is sequential, and, with the reducing batch window that most of you are experiencing, any improvement in the handling of sequential data is welcome. The RAMAC 3 drawer as well as the RAMAC and RAMAC 2 drawer have enhancements in their design to improve processing for sequential update writes. This improvement is in addition to the faster data rate that the high-speed device adapter provides on the RAMAC 3 drawer.

RAMAC 3 Drawer Capacity

IBM
San Jose
© IBM Corporation 1996



© IBM Corporation 1996

RAMAC 3 Drawer Capacity

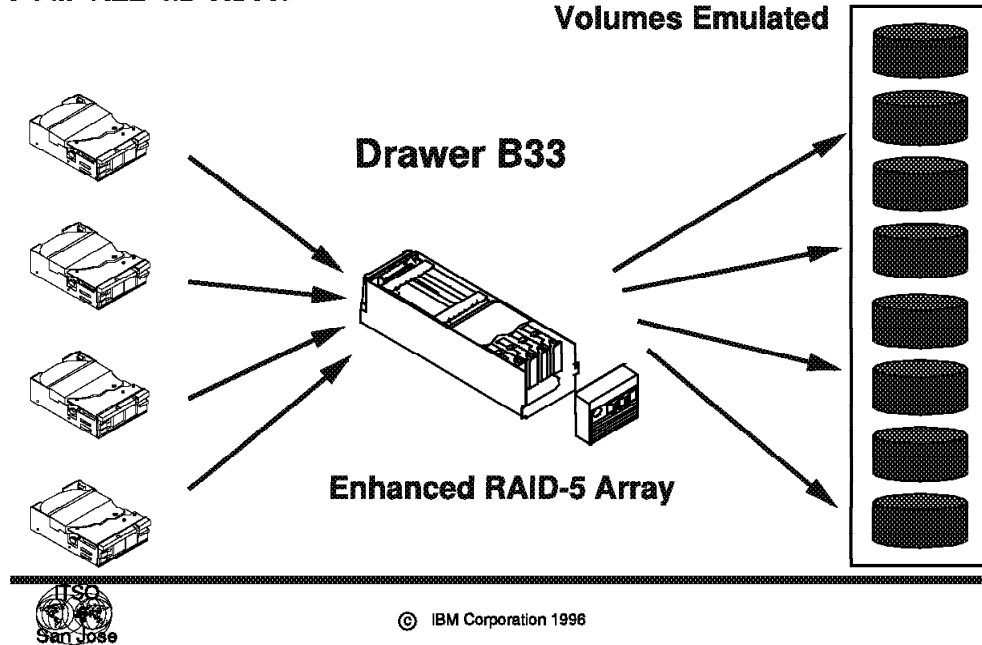
A RAMAC 3 drawer holds 22.7 GB of data in 3390-3 format. If this capacity is compared with the capacity of a 3390-A38, a box that is still installed in many installations today, you can see the dramatic reduction in the physical space needed to store your data. A full frame of RAMAC 3 drawers containing 363.2 GB is equivalent to four full strings of 3390-3, or 12 boxes.

RAMAC 3 Drawer Logical Volume Emulation



Four 9.22 GB HDAs

Eight Logical 3390-3
Volumes Emulated



RAMAC 3 Drawer Logical Volume Emulation

The additional capacity in the RAMAC 3 Array Storage is delivered by an increase in the number of logical volumes that can be accessed in a RAMAC 3 drawer while retaining an optimal physical array size of four disk drives. The RAMAC 3 drawer emulates eight 3390-3 logical volumes that can be installed in either 3390-3 track format or 3380 track format. The RAMAC 3 drawer always emulates 3390-3 track geometry, with the storage control (3990-6 or 9390) providing the 3380 track format function in addition.

RAMAC Drawer Logical Volume Mapping



DISK1	DISK2	DISK3	DISK4
V0C0T0	V0C0T1	V0C0T2	PARITY
V0C0T3	V0C0T4	V0C0T5	PARITY
V0C0T6	V0C0T7	V0C0T8	PARITY
V0C0T9	V0C0T10	V0C0T11	PARITY
V0C0T12	V0C0T13	V0C0T14	PARITY
V1C0T0	V1C0T1	V1C0T2	PARITY
V1C0T3	V1C0T4	V1C0T5	PARITY
V1C0T6	V1C0T7	V1C0T8	PARITY
V1C0T9	V1C0T10	V1C0T11	PARITY
V1C0T12	V1C0T13	V1C0T14	PARITY
V0C1T0	V0C1T1	PARITY	V0C1T2
V0C1T3	V0C1T4	PARITY	V0C1T5
V0C1T6	V0C1T7	PARITY	V0C1T8
V0C1T9	V0C1T10	PARITY	V0C1T11
V0C1T12	V0C1T13	PARITY	V0C1T14
V1C1T0	V1C1T1	PARITY	V1C1T2
V1C1T3	V1C1T4	PARITY	V1C1T5
V1C1T6	V1C1T7	PARITY	V1C1T8
V1C1T9	V1C1T10	PARITY	V1C1T11
V1C1T12	V1C1T13	PARITY	V1C1T14



© IBM Corporation 1996

RAMAC Drawer Logical Volume Mapping

The four disk drives in the RAMAC drawer use fixed-block architecture (FBA) to enable RAID-5 distributed parity. However, system control programs (SCPs) such as MVS mandate count key data (CKD) channel architecture and disk format. Therefore an emulation process is needed that translates CKD format to FBA format and vice versa. This emulation process is performed in the RAMAC drawer. The fixed-block disks in the RAMAC drawer are formatted with a number of 688-byte SCSI sectors or blocks. Eighty-nine SCSI blocks are needed to emulate a CKD track. The CKD data can be striped across the four disks in the array in a number of ways, and careful consideration was given to this aspect of the drawer design.

The foil shows how the RAMAC drawer maps the CKD data to the four disk drives in the RAMAC drawer array. Remember that the RAMAC drawer emulates two 3390-3 logical volumes—call them logical volume 0 and logical volume 1. Notice that the striping increment is one CKD logical track. The foil shows V0C0T0 as the first data stripe written to disk 1—this refers to logical volume 0, logical cylinder 0, logical track 0 of a 3390-3 volume. This logical track is 57 KB long and occupies 89 SCSI sectors on disk 1 in the array. Similarly, V0C0T1 (volume 0, cylinder 0, track 1) is written to the first 89 SCSI blocks of disk 2 in the array, and V0C0T2 (volume 0, cylinder 0, track 2) is written to the first 89 SCSI sectors of disk 3 in the array. The first 89 SCSI sectors of disk 4 contain the parity information for these three CKD data tracks. The second 89 SCSI sectors of disk 1 contain the data for 3390-3 logical volume 0, cylinder 0, track 3 (V0C0T3).

RAID-5 architecture is characterized by what is referred to as *distributed parity*; that is, the parity information that is used to rebuild the contents of any failed disk drive is spread across all of the disk drives to avoid “single-drive contention” during parity processing. Notice how the RAMAC drawer implements distributed parity. Parity is not rotated among the disk drives in a RAMAC drawer array until an entire logical cylinder from each logical volume is written. The foil shows that the 15 logical tracks in the first logical cylinder are written for each logical volume to disks 1, 2, and 3 before parity rotates from disk 4 to disk 3. The fifteenth track in the first logical cylinder of volume 0 is V0C0T14 and is written to disk 3 in the array, and the fifteenth track in the first logical cylinder of volume 1 is V1C0T14 and is also written to disk 3 in the array.

The logical tracks that are emulated in a two-volume RAMAC drawer will always occupy these locations on the disk drives under normal circumstances. This locality of data offers a significant benefit especially for sequential workloads because the write penalty normally associated with RAID-5 architecture is avoided. The RAMAC drawer logical volume mapping is therefore instrumental in avoiding the performance drawback inherent in RAID-5 devices. Write penalty avoidance is described in more detail in Appendix A.

RAMAC 2 Drawer Logical Volume Mapping

DISK1	DISK2	DISK3	DISK4
V0C0T0	V0C0T1	V0C0T2	PARITY
V0C0T12	V0C0T13	V0C0T14	PARITY
V1C0T0	V1C0T1	V1C0T2	PARITY
V1C0T12	V1C0T13	V1C0T14	PARITY
V2C0T0	V2C0T1	V2C0T2	PARITY
V2C0T12	V2C0T13	V2C0T14	PARITY
V3C0T0	V3C0T1	V3C0T2	PARITY
V3C0T12	V3C0T13	V3C0T14	PARITY
V0C1T0	V0C1T1	PARITY	V0C1T2
V0C1T12	V0C1T13	PARITY	V0C1T14
V1C1T0	V1C1T1	PARITY	V1C1T2
V1C1T12	V1C1T13	PARITY	V1C1T14



© IBM Corporation 1996

RAMAC 2 Drawer Logical Volume Mapping

With the arrival of RAMAC 2, the capacity of the drawer is doubled to 11.35 GB. However, the concept of logical volume emulation is preserved, and the increase in disk capacity is realized through a doubling of addressable volumes as opposed to higher-capacity volumes, for example. How does this affect the mapping scheme described in the previous foil? Can the same design objective of write-penalty avoidance be achieved for sequential workloads?

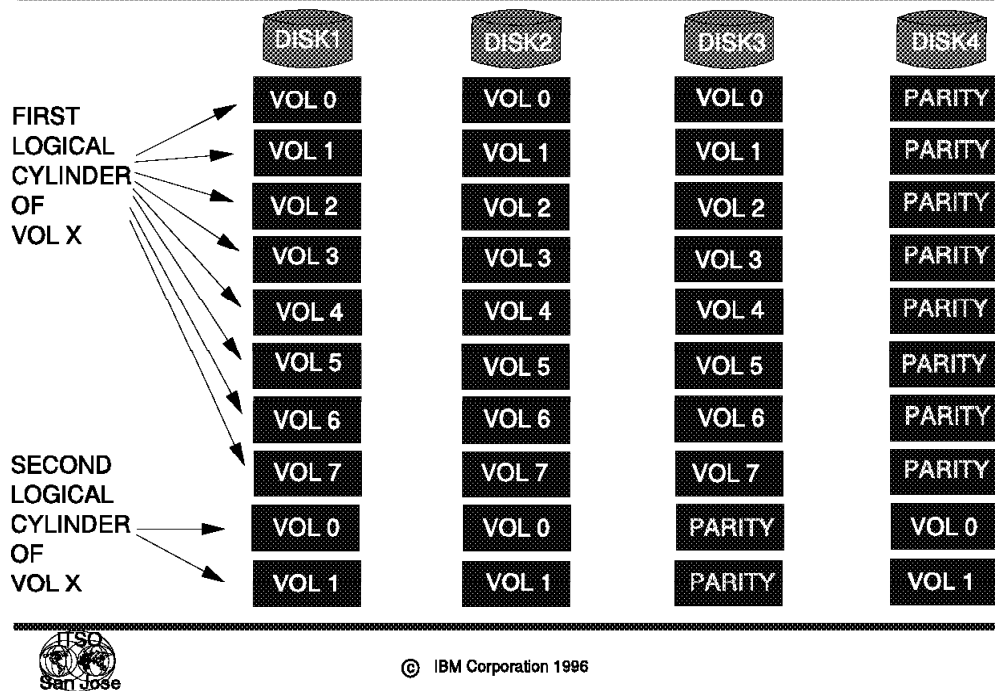
The answer is yes. The foil shows the mapping details of the RAMAC 2 drawer that is emulating four logical volumes as opposed to two logical volumes in the RAMAC drawer. Notice that the same concept of supporting RAID-5 distributed parity is retained by rotating the parity information across all of the disks in the array. The only change between RAMAC and RAMAC 2 is the amount of data written to the disks in the array before the parity is rotated. Recall from the RAMAC drawer mapping design that an entire cylinder of data (15 logical tracks) is written for each volume being emulated in the drawer before parity is rotated. The same technique is used in the RAMAC 2 drawer.

Notice that the parity information for the first cylinder of volumes 0, 1, 2, and 3 is stored on disk 4. The range of tracks for the first cylinder of each of the four volumes is:

Volume 0 - V0C0T0---V0C0T14
 Volume 1 - V1C0T0---V1C0T14
 Volume 2 - V2C0T0---V2C0T14
 Volume 3 - V3C0T0---V3C0T14

Notice that all of these tracks are written to disks 1, 2, and 3 in the RAMAC 2 drawer array. The second cylinder of logical volumes 0, 1, 2, and 3 will be written to disks 1, 2, and 4 with corresponding parity written to disk 3. Therefore distributed parity is maintained in the RAMAC 2 array, and the adjacency of data is preserved to optimize sequential performance while increasing drawer capacity.

RAMAC 3 Drawer Logical Volume Mapping



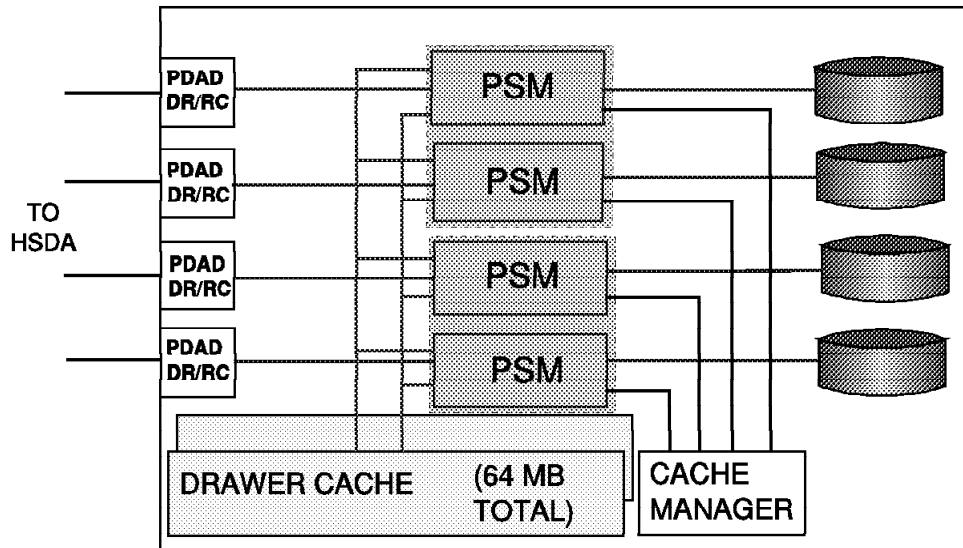
RAMAC 3 Drawer Logical Volume Mapping

It should come as no surprise to discover that the drawer mapping design for RAMAC 3 uses the same methods of the predecessor products while keeping the array size optimized at four drives and yet doubling capacity for the second time to 22.7 GB. The RAMAC 3 drawer emulates eight logical 3390-3 volumes, installed in either 3390-3 or 3380 track format. The distribution of parity in the RAMAC 3 array follows the same rules as before—write out one logical cylinder (15 logical tracks) of each logical volume being emulated, then rotate parity. Again, this technique increases capacity through increased device addressability and enables the enhanced RAID-5 design of the RAMAC drawer to take advantage of the locality of CKD track data to provide significant benefits for sequential workloads. Refer to Appendix A for details on how the RAMAC Array Family drawer technology and design avoid RAID-5 write penalty performance for sequential I/O.

RAMAC 3 Array Drawer Technology



RAMAC 3 ARRAY DRAWER



© IBM Corporation 1996

RAMAC 3 Array Drawer Technology

The RAMAC 3 drawer incorporates the same technology and design as the RAMAC and RAMAC 2 drawers. Improvements have been made to the drawers as each generation of the RAMAC Array Family addresses the need for increased capacity, higher availability, better reliability, improved data integrity, and enhanced performance in more compact and cost-effective configurations.

The RAMAC, RAMAC 2, and RAMAC 3 drawers use a "multiengine" approach to deliver the function required of an enhanced RAID-5 device. Recognizing the processing requirements of CKD-to-FBA translation, CU-device protocol conversion, parity computation, destage management, data transfer, and maintenance and recovery, the RAMAC, RAMAC 2, and RAMAC 3 drawers exploit a highly parallel design that overlaps execution of many tasks in an advanced RAID architecture subsystem. Like the RAMAC and RAMAC 2 drawers, the RAMAC 3 drawer implements a fault-tolerant design based around the port SCSI manager (PSM) component

There are two PSM cards in each drawer, and for the RAMAC 3 drawer the PSM is the third generation card. Each PSM card contains two path managers that are responsible for data transfer between the disk and the drawer cache and the drawer cache and the channel. The PSM also performs other functions that are described on the next foil.

Each PSM card has two paths to drawer cache, providing a further level of fault tolerance at the pathing level and a connection to the cache manager used for distributed communications and recovery. The PSMs can perform a takeover

function if the cache manager is not available—this is a different kind of fault tolerant design that avoids component replication but still delivers automatic recovery and application data access.

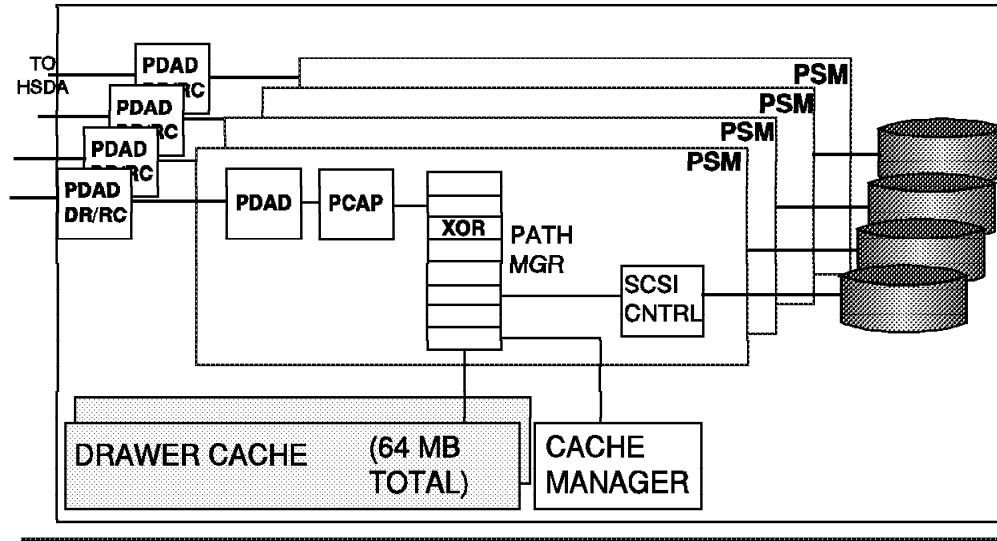
The PSMs connect to the PDAD driver and receiver components, which in turn connect to the high-speed device adapters (HSDAs) in the RAMAC 3 Storage Frame. The HSDAs in turn connect to the DDC adapters in the 3990-6 and the 9390. At the lower interface level, the PSMs also connect to the new fixed-block Ultrastar disk drives that operate at data rates of up to 15.4 MB/s. The SCSI interface between the PSM and the disk is a fast-and-wide single-ended interface supporting data rates of up to 20 MB/s.

The next foil takes an in-depth look at the RAMAC drawer and describes the changes to the drawer components incorporated into the RAMAC 3 Array Storage.

RAMAC 3 Array Drawer Technology...



RAMAC 3 ARRAY DRAWER



© IBM Corporation 1996

RAMAC 3 Array Drawer Technology

The PSM card in the RAMAC, RAMAC 2, and RAMAC 3 drawers contain the "multiengine" hub of RAMAC drawer operations. The nature of the work processed by the RAMAC drawer lends itself to highly parallel operations—for example, an upper interface data transfer, in process simultaneously with a lower interface staging operation, in process simultaneously with a RAID-5 parity computation. Overlapping these tasks by executing them simultaneously on multiple engines or microprocessors results in the fastest of RAID-5 implementations.

The PSM contains several key components that "join forces" to provide the highly parallel, fast execution, advanced RAID-5 RAMAC operation.

The path manager function of the PSM enables connectivity and data communications and transfer between the interfaces. The path manager is the "gateway" to drawer cache and uses fault-tolerant pathing. It "links" the upper drawer interface with the lower drawer interface through intelligent port-management.

The path manager component of PSM connects to the PCAP processor that provides the CKD-to-FBA emulation needed as data flows from the 9390 or 3990-6 to the Ultrastar 2XP disk drives. The PCAP processor connects to the PDAD that interfaces to the control unit and is used to transfer host data to and from drawer cache.

The path manager also connects to the SCSI controller and facilitates data transfer between the drawer cache and Ultrastar 2XP disk drives. The SCSI controller connects to the disk drives through disk cache and uses a 20 MB/s fast-and-wide single-ended interface.

The RAID-5 parity processing function is performed in the hardware by the path manager component and uses an exclusive-or (XOR) Boolean operation that computes a binary bit pattern from the three data tracks or from "old" and new versions of updated data and existing parity.

Drawer cache is used to hold application data and assist with RAID-5 operations, advanced destage management, and logical volume mapping. The drawer cache is fault tolerant; that is, it is contained on separate cards, and it is nonvolatile because a battery provides an alternative source of power that can sustain the contents of the cache if there is a room power outage. The new storage control NVS management changes use the drawer cache as a nonvolatile store instead of using the storage control NVS (see "Enhanced NVS Management" on page 95). In this case, the data resides in both controller cache and drawer cache. In the case of a power failure, the drawer ensures that any modified data in the drawer cache is destaged successfully to disk. The drawer stages tracks or partial tracks from the disk drives, so unproductive staging is avoided if a reference is made to a block that is not located at the beginning of a track. The drawer stages tracks or records from drawer cache to the storage control depending on the mode of access—track or record level caching.

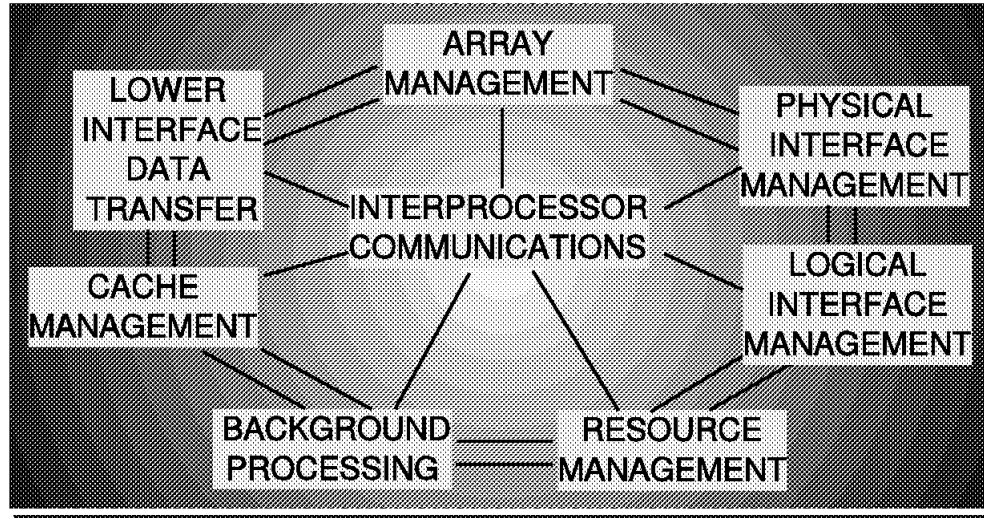
The cache manager microprocessor manages drawer cache contents and schedules the processing of parity in support of the RAID-5 operation. Cache manager outage does not result in data loss or data access loss. If the cache manager malfunctions, Dynamic Sparing (if configured) will start copying the data to the spare drawer, and the PSMs will load the cache manager microcode and take over the function that allows drawer operation to continue.

Within the RAMAC 3 drawer, there are both hardware and microcode enhancements. The major drawer performance enhancement is the replacement of the device adapter of RAMAC and RAMAC 2 with a higher-performance PDAD to support the higher data rate to and from the storage control. The new adapter, a third generation chip, supports a data rate of up to 11.9 MB/s.

RAMAC Drawer - System Components



THE WEB OF ADVANCED RAID-5 FUNCTION



© IBM Corporation 1996

RAMAC Drawer - System Components

The multiengine hardware design of the RAMAC drawer enables multiple I/O and RAID-5 tasks to be performed in parallel. Overlapping RAMAC device tasks results in the highest levels of RAID-5 data redundancy while realizing high performance goals. To meet the aggressive goal of delivering the highest levels of data availability with the highest levels of subsystem performance, the hardware design of the RAMAC drawer is complemented by a multitasking microcode operating system—the highly parallel operation of the RAMAC drawer is a result of this operating system exploiting the multiengine hardware design. The execution of the RAMAC drawer tasks are transparent to application code.

The components of the RAMAC drawer's operating system are distributed among the microprocessors in the drawer for maximum efficiency.

Array management

The RAMAC drawer's array management tasks are coordinated by a supervisor component that commands and regulates the initiation of the many tasks that are processed in the drawer. Although these tasks are executed by other drawer components, the array manager supervises their scheduling. For example, the array manager initiates and supervises the scheduling of parity computations and the staging and destaging of data. The coordination of the many processes in an advanced RAID-5 data storage device is vital if the product is to evolve from a stable base. The RAMAC 3 Array drawer is an example of such an evolution.

Physical interface management

The RAMAC Array family of products (excluding the RAMAC Virtual Array, Scalable Array, and Electronic Array) implements a multilevel cache structure designed to provide high-performance RAID-5 capability. The RAMAC Array drawer recognizes the special needs of data transfer operations in a multilevel cache design and dedicates drawer hardware and microcode resources to ensure that data transfer is optimized. The RAMAC Array drawer's physical interface management system controls the translation of data as it flows to and from the storage control over different interfaces and uses different protocols. The RAMAC Array drawer executes the physical interface processing required to send and receive data to and from the host.

Lower interface data transfer management

The RAMAC Array drawer's lower interface data transfer management system is responsible for transferring data from the disk drives to the drawer cache. Having dedicated drawer elements performing data staging and destaging enables other drawer processes to execute in parallel and under the supervision of the array management system.

Interprocessor communications

The RAMAC Array drawer's multiengine design enables the components of the operating system to execute in parallel in addition to multitasking. The signaling and command processing that enable control to be passed among the various drawer components are managed by the interprocessor communications system. The interprocessor communications component in the drawer ensures that important I/O and RAID-5 tasks are dispatched properly across multiple processor platforms.

Cache management

The storage control reads data from and writes it to drawer cache, and the drawer reads and writes data to and from the disk drives. The drawer cache is at the center of the upper and lower interface bidirectional data transfers, and it must provide data integrity, high availability, and high-performance RAID-5 operations support. The drawer's cache management system ensures the validity of cache contents as data is staged, destaged, promoted, and demoted. The cache management system works with the specialized hardware engines in the drawer to maintain cache segment status for operations such as parity calculations, and padding (to end of track).

Logical interface management

The drawer executes the logical interface processing that makes data communications meaningful by interpreting channel and drawer commands. The host system control program (SCP) reads and writes data in count-key-data (CKD) format. The RAMAC Array drawer enables cost-effective RAID-5 operations by storing data on Ultrastar 2XP fixed block disk drives. The conversion between the two disk formats is performed by the PCAP processor in the drawer.

Background processing

All RAID-5 data storage subsystems must perform the tasks that enable them to qualify for RAID-5 designation. However, the RAMAC Array drawer changes the rulebook on how RAID-5 devices should behave. One of the many examples of how the RAMAC Array drawer expands the frontiers of RAID-5 operation is demonstrated by the activities performed by the drawer's background management procedures. The background processing enables functions like disk scrubbing, data regeneration, and advanced destage management to execute on a nonintrusive basis with I/O and RAID-5 operations. The tasks that lend themselves to background execution deliver important availability, data integrity, and performance functions. Background disk scrubbing enables data sector errors to be detected and corrected before applications make requests for the data. Background data regeneration enables application data to be reconstructed on newly replaced disk drives, thus delivering the highest levels of availability by restoring RAID-5 capability. Advanced destage management support ensures that write penalty avoidance opportunities are detected and acted on to improve drawer resource utilization and improve performance.

Resource management

The effectiveness and efficiency of the RAMAC Array drawer's RAID-5 operations are delivered through a combination of high-speed technology, highly parallel design, and advanced function. The RAMAC Array drawer ensures that drawer components are shared optimally, and that data integrity is not compromised when multiple accesses to a single drawer or data resources are requested. The drawer accomplishes these tasks through its resource management system. The resource management system allocates drawer resources such as drawer cache segments and control memory according to direction from the array management system. The resource management system also provides locking functions that enable access to and prevent corruption of application data when multiple requests are issued for that resource—for example, access to a track image residing in drawer cache is coordinated by the drawer's resource management system.

— RAMAC Array drawer - RAID-5 leadership —

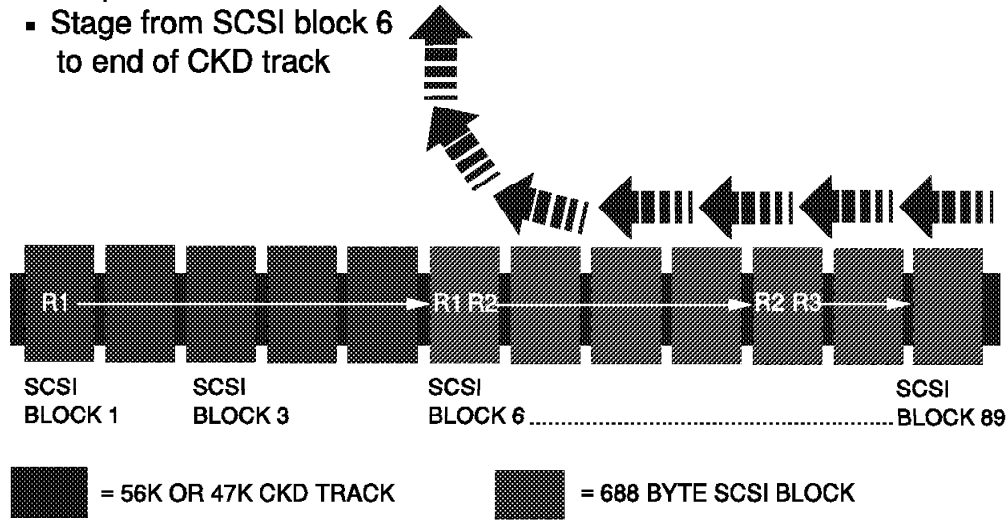
Highly parallel, multiengine, multitasking, distributed RAID-5 data storage

RAMAC 3 Drawer Hardware Assists



CKD/SCSI MAPPING

- Request R2
- Stage from SCSI block 6 to end of CKD track



© IBM Corporation 1996

RAMAC 3 Drawer Hardware Assists

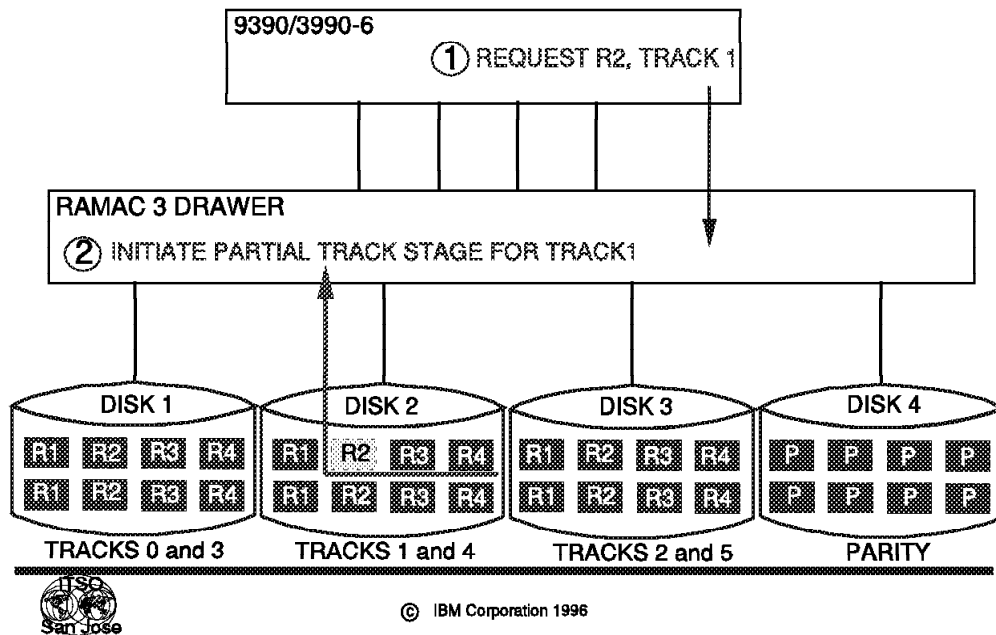
This and the next two foils describe the hardware-assisted function that improves RAMAC 3 data transfer operations. In the RAMAC and RAMAC 2 drawer designs, the data transfer setup operations described in the next three foils are executed by drawer microcode.

RAMAC, RAMAC 2, and RAMAC 3 implement efficient staging and destaging operations from and to the fixed block disk drives. The unit of staging from the disk drives to the drawer cache is the partial CKD track. The unit of destaging from the drawer cache to the disk drives is the SCSI sector (688 bytes). The drawer maintains the relationship between the CKD view of the data and the FBA view of the data. When a record is requested, the drawer and the disk drive cooperate to stage in the SCSI sectors containing the record and the SCSI sectors containing the data on the remainder of the track. In the example on the foil, a request for R2 is made. A part of R2 shares the same SCSI sector as a part of R1, so the drawer stages from SCSI block 6 to the end of the CKD track (SCSI block 89). R3 is also staged because it resides on the same track—this is called *partial track staging*. The next two foils show how the RAMAC 3 Array drawer uses a hardware-assisted function to improve partial track staging performance.

RAMAC 3 Drawer Hardware Assists...



DATA TRANSFER/STAGING OVERLAP



RAMAC 3 Drawer Hardware Assists...

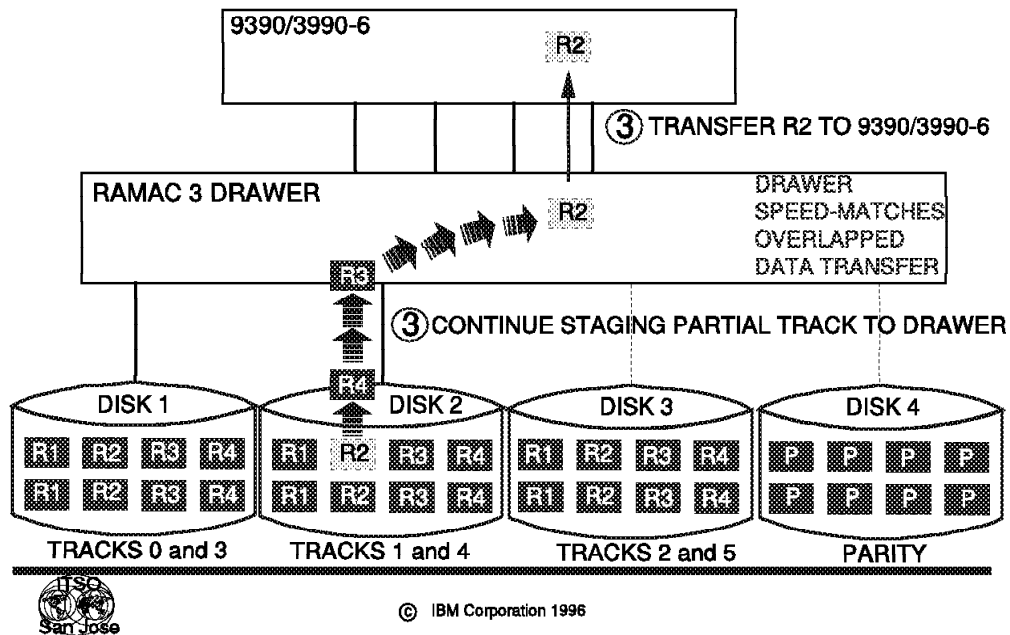
The foil shows how the RAMAC, RAMAC 2, and RAMAC 3 drawers optimize the data transfer from disk drive to drawer cache to control unit cache. This "journey" through the subsystem hierarchy is enhanced by overlapping the staging of requested data from disk to drawer cache with the staging of data from drawer cache to control unit cache.

When the control unit requests a record that is not in drawer cache, the drawer initiates the staging of the record from the disk. The foil describes a request for record 2 on logical track 1 (this record is stored on disk drive 2). In anticipation of usage of adjacent records, the drawer stages from the record requested to the end of the logical track—that is, records R2, R3, and R4 from track one on disk 2. At this point, steps 1 and 2 have been executed. The next foil shows the subsequent steps that the drawer takes in the partial track staging operation.

RAMAC 3 Drawer Hardware Assists...



DATA TRANSFER/STAGING OVERLAP



RAMAC 3 Drawer Hardware Assists...

Step 3 involves staging R2, R3, and R4 from the disk to the drawer cache. The RAMAC, RAMAC 2, and RAMAC 3 drawers do not wait for the entire partial track to be staged, however. Instead, when the requested record is in the drawer cache and available for transfer to the control unit, the drawer signals readiness to the host. The drawer and the control unit then transfer the record while the remaining records in the track are simultaneously staged from this disk drive to the drawer cache. This process offers a fast record staging technique in the event of subsystem and drawer cache miss.

The setup for this staging overlap is managed in the microcode in the RAMAC and RAMAC 2 drawers but is implemented in the hardware in the RAMAC 3 drawer. This change results in faster data prefetch processing for RAMAC 3 Array Storage.

RAMAC Data Integrity and Availability



- Disk scrubbing
 - RAMAC 3, RAMAC 2, and RAMAC
 - Drawer logic reads all customer data
 - 7-day cycle
 - Low-priority background task
 - If bad block, use RAID-5 to recalculate data and rewrite to spare block
 - Recalculates and rewrites parity on monthly cycle

Provides Improved Data Integrity and Availability



© IBM Corporation 1996

RAMAC Data Integrity and Availability

Several enhancements have been made to data integrity and availability.

Disk scrubbing

Disk scrubbing, available on RAMAC, RAMAC 2, and RAMAC 3, is designed to provide a continuous check of the disks and your data. The drawer logic reads all customer data on all four disks in the array on a 7-day cycle. This activity takes place as a background activity when the drawer is idle. If a bad block is encountered, the drawer uses its RAID-5 capability to recalculate the bad data and rewrite it to a spare block. This is a joint activity between the disk and the drawer.

The benefits of these changes are increased data integrity and availability.

RAMAC 3 Improved Data Integrity



- SCSI header checking on every block
 - Implemented in hardware
 - ▶ *Facilitated by CMOS packaging*
 - Circumvents tradeoff between performance and data integrity
 - ▶ *Previous implementations performed in microcode*

**Provides Maximum Data Integrity Protection
without Impacting Performance**



© IBM Corporation 1996

RAMAC 3 Improved Data Integrity

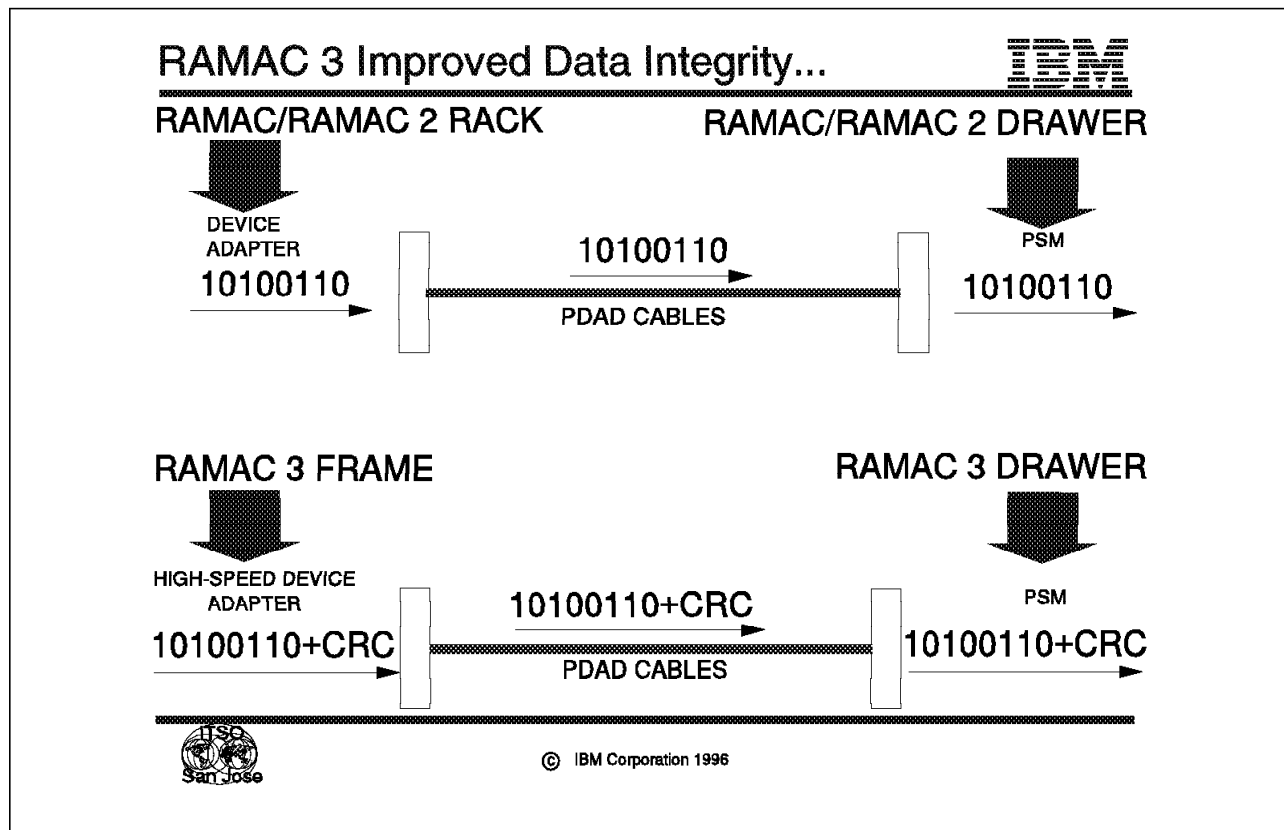
The third generation of the RAMAC Array Family, RAMAC 3, further improves the data integrity of the product line by capitalizing on CMOS technology.

SCSI header checking on every block

RAMAC 3 implements an improvement to the SCSI block header checking technique by checking the SCSI header on every block read, ensuring that even the remotest possibility of reading the wrong block is eliminated.

Checking the SCSI block header on every block is possible because the RAMAC 3 drawer implements the process in the hardware. CMOS technology in the RAMAC 3 drawer has enabled the additional function to be incorporated into the existing circuitry without requiring additional space on the cards. RAMAC 3 CMOS packaging is thus the enabler of improved data integrity and availability.

By checking every header block, the likelihood of having to invoke other drawer recovery functions is avoided, and there is no impact on performance because the task is now performed in the RAMAC 3 Array drawer hardware.



RAMAC 3 Improved Data Integrity...

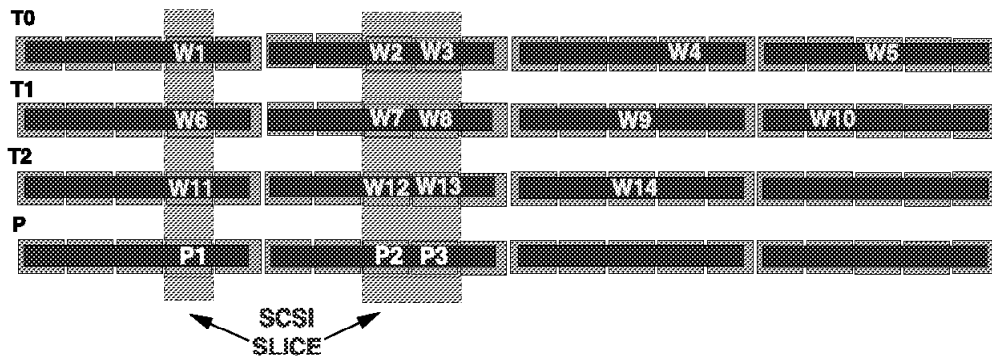
A chain is only as strong as its weakest link. The RAMAC 3 drawer delivers additional levels of error detection and correction that apply to the most innocuous of storage subsystem components—cables, drivers, and receivers. With RAMAC and RAMAC 2, drawer recovery measures such as RAID-5 operations are used to protect the integrity of the data as it flows to and from the storage control.

In RAMAC 3 drawers a cyclic redundancy check (CRC) has been incorporated into the interface logic to provide a fast detection and correction technique for data transferred from the HSDA in the RAMAC 3 Storage Frame to the PDAD in the RAMAC 3 drawer. The CRC across the PDAD interface enables errors in cabling, drivers, and receivers to be detected and corrected as rapidly as possible without having to invoke other drawer recovery procedures at a later stage of the I/O process. Recovery avoidance without compromising data integrity and availability ensures faster, smoother data transfers.

Improved Sequential Write Processing



Improved drawer processing for sequential update writes



- Drawer "scans" for updates in same SCSI block slice
- Performs partial track destage on SCSI block boundaries
- No RAID-5 write penalty

= SCSI BLOCK

= CKD RECORD



© IBM Corporation 1996

Improved Sequential Write Processing

Algorithm changes have been introduced to the RAMAC 3 drawer to improve the drawer destage performance for some write scenarios.

The foil shows an update-write pattern to three CKD tracks that occupy corresponding positions on three separate disk drives in the RAMAC 3 RAID-5 array. T0, T1, and T2 are the three data tracks, and P is the parity track. Parity is calculated by XORing the three data tracks to generate the parity track. Because the data on these tracks is located in corresponding locations on different disks, the tracks are considered to be in the same "slice". A slice can be considered as all of the data needed to calculate parity for that part of the array, so updates W1 through W14 are written to tracks that are in the same slice.

The RAMAC 3 disks use a fixed block architecture, therefore, the data on each track is written in fixed segments or blocks as shown on the foil (for example, W1 is contained in a fixed, 688-byte block). If corresponding blocks in the same slice are updated, the RAMAC, RAMAC 2, and RAMAC 3 drawers can calculate the parity for that "partial track slice," and only the changed data is prepared for destage—not the whole track. Referring to the foil, we see that W1, W6, and W11 are corresponding blocks in the same partial track slice (let's call it a SCSI block slice). The slice is indicated on the foil by the vertical bar shadowing the W1, W6, and W11 write updates. Similarly, W2 and W4 are in different SCSI block slices, but W9 and W14 are in the same SCSI block slice (however, this slice is incomplete).

The RAMAC, RAMAC 2, and RAMAC 3 drawers execute a function that scans the drawer cache contents to determine whether update writes have completed to SCSI blocks on different tracks that are in the same track slice. In addition, a further level of scanning takes place to determine whether update writes are targeted to all of the blocks in the same SCSI block slice (W1, W6, and W11). When the drawer detects such a reference pattern, it calculates the parity for the three SCSI data blocks without incurring the write penalty, and only the changed data (W1, W6, W11, and P1) has to be “scheduled” for writing to the disks.

The algorithm change applies to update writes. Typically an update write would have a RAID-5 write penalty because only part of the data on one track is updated. However, some update data is in fact sequential in nature—and updates records sequentially. A classic example is DB2 logging, where the log data set is preformatted, and all of the write commands are update writes.

The advanced drawer cache scanning logic enables RAMAC to recognize data that has been changed across all tracks of a slice and across all blocks within the slice and is performed as a background task independent of production I/O processing. This logic is similar to typical RAMAC sequential writing where the RAMAC Array operates like a RAID-3 array, but in this case only the changed blocks on each track have to be written. The enhanced drawer logic eliminates the RAID-5 write penalty for some update write activity.

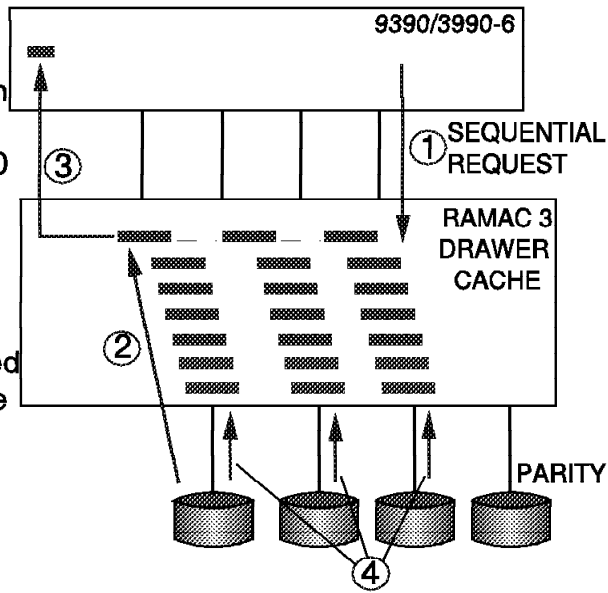
Although DB2 logging is a good example of where the enhanced sequential algorithm can be used, the drawer logic will look for any combination of data in the drawer cache that can be written in this way.

Appendix A describes the RAMAC drawer’s advanced destage management processes in more detail.

RAMAC 3 Drawer Sequential Prestaging



- Requested track and all tracks to end of cylinder staged to drawer cache on detection of sequential request from 9390 or 3990
 - Triggered by sequential indicator in DX or by
 - Sequential detect by storage control
- Additional six tracks staged to "keep ahead" of storage control as data is transferred
- Drawer speed-matches overlapped data transfer



© IBM Corporation 1996

RAMAC Drawer Sequential Prestaging

A large part of the data that most customers process is sequential. Any improvements in the design and algorithms that help sequential throughput help reduce batch elapsed times. The improvements to drawer sequential staging, also available on existing RAMAC and RAMAC 2 drawers, helps improve the drawer handling of sequential data requests from the storage control.

Requested track and all tracks to end of cylinder...

On a request from the RAMAC 3 or 3990-6 storage control that sequential data is required, the drawer starts staging up the requested track, and the remaining tracks in the cylinder. If data from track 0 is requested, the drawer will stage all of the records from the record referenced to the end of the track, followed by all of the records on tracks 1 through 14. Tracks are staged from disk to drawer cache in parallel.

The RAMAC 3 or 3990-6 storage control determines that sequential processing is required through either the sequential indicator in the Define Extent command or by detecting that the data being read is sequential. See foil "Sequential Detect" on page 93 for more details on the sequential detection algorithm.

Additional six tracks staged to "keep ahead"...

As the data is transferred from the drawer to the RAMAC 3 or 3990-6 storage control, the drawer continues to read from its disks, maintaining six tracks ahead of the storage control.

Drawer speed-matches overlapped data transfer

The RAMAC 3 Array Storage subsystem uses a data transfer hierarchy that involves disk drives, drawer cache, and subsystem cache. With the new HSDA, enhanced PDAD, and high-speed cables, data transfers are now faster than ever in a RAMAC environment. In addition, the RAMAC 3 drawer and the 9390 and 3990-6 storage use techniques to optimize I/O operations by overlapping data transfer. The RAMAC 3 drawer electronics provide speed-matching capability between the disk drives and the storage control to ensure that data flows smoothly over upper and lower interfaces that use components with different speed capabilities.

IBM RAMAC 3 Drawer - Summary



HIGH PERFORMANCE

- Highly parallel RAID-5 operations
 - Multiengine hardware
 - Multitasking microcode
- Faster PDAD
- Faster h/w-assisted data prefetch
- Sequential processing
 - Sequential detect support
 - Advanced destage management
- Logical volume mapping
- Ultrastar 2XP disk drive

HIGH CAPACITY

- Ultrastar 2XP disk drives
- Eight logical volume emulation support

DATA INTEGRITY

- H/W-assisted SCSI block checking
- CRC across CU-drawer interface

HIGH AVAILABILITY

- Third generation RAID-5
- Fault-tolerant design



© IBM Corporation 1996

IBM RAMAC 3 Drawer - Summary

The foil summarizes the improvements made to the RAMAC 3 drawer in the categories of high performance, high capacity, data integrity, and high availability.

Highly parallel RAID-5 operations

The RAMAC Array Family provides an enhanced RAID-5 storage device that eliminates many of the drawbacks associated with RAID-5 architecture. The highly parallel operations of the RAMAC 3 drawer are delivered through a multiengine hardware design and sophisticated drawer operating system. Together, the multiengine hardware and operating system overlap RAID-5 and I/O processing activities to such a degree that drawer component utilization and application service times are optimized across a wide range of subsystem activities.

Faster PDAD

The faster PDAD in the RAMAC 3 drawer supports the higher data rates now achievable on the storage subsystem lower interface. The high-speed PDAD located in the drawer complements the HSDA located in the Storage Frame and the high-speed DDC cables. This "high-speed trio" of components results in faster data transfers that reduce response times and increase throughput for sequential workloads.

Faster h/w-assisted prefetch

The RAMAC 3 drawer implements faster data prefetch from disk drive to drawer cache to subsystem cache. The faster data prefetch is achieved by using drawer hardware to assist in preparing for data exchanges between the drawer and the storage control.

Sequential detect support

The RAMAC 3 drawer prestages tracks in anticipation of use by sequential applications. Sequential prestaging is exploited further when the storage control signals sequential intent for workloads that traditionally do not set the sequential access indicator in the channel program—sequential detect. The prestaging of tracks to drawer cache before the application requests them enables fast transfer to the storage control using the high-speed PDAD, DDC cables, and HSDA. Thus, cache access rather than disk access satisfies the data transfer requirement.

Advanced destage management

The RAMAC 3 drawer uses advanced destage management techniques to optimize RAID-5 operations and reduce disk usage for asynchronous writes to the disk drives. These algorithms enable the upper interface to concentrate on providing optimum response times, and the lower interface to concentrate on supporting maximum throughput.

Logical volume mapping

The conversion of CKD to fixed block disk format must be accompanied by a method that enables efficient storage and retrieval of application data. In addition, application performance is sometimes dependent on locality of reference of successive data requests. RAMAC 3 Array Storage supports efficient storage and retrieval of data that has been converted from CKD format to FB format. This is accomplished with the logical volume mapping techniques used by RAMAC 3 drawers. The logical volume mapping techniques permit CKD-to-FB conversion at the same time as preserving the locality of reference that is important for sequential I/O performance.

High-performance Ultrastar 2XP disk drive

The RAMAC 3 drawer uses the Ultrastar 2XP disk drives for RAID-5 storage. The Ultrastar 2XP contributes to the RAMAC 3 drawer performance through faster rotational speed (7200 rpm), faster media data rate (10.3-15.4 MB/s), and enhancements to the disk architecture such as No-ID sector format. The Ultrastar 2XP disk drive is discussed in more detail in Chapter 4.

High-capacity Ultrastar 2XP disk drive

RAMAC 3 uses the Ultrastar 2XP 9.22 GB disk drive in its RAID-5 array. Use of the Ultrastar 2XP has doubled capacity when compared to the RAMAC 2 Array DASD.

Eight logical volume emulation support

The capacity improvements are achieved in a four-disk array through specialized logical mapping techniques that support increased device addressability.

H/W-assisted SCSI block checking

The RAMAC 3 drawer incorporates the highest levels of data integrity checking in the industry and takes advantage of CMOS packaging to broaden the scope of error detection and correction. The RAMAC 3 drawer now checks every SCSI block for bit-pattern errors that threaten data integrity.

CRC across CU-drawer interface

When data flows through a hierarchy of components in a highly parallel, distributed RAID-5 design, great care must be taken to protect the validity of application data. The RAMAC 3 Array drawer is “obsessive” about the integrity of customer’s data. Cyclic redundancy checks now determine the integrity of data when it is received at the driver-receiver drawer interface. CRC processing at the driver-receiver interface enables detection and correction of errors caused by driver-receiver hardware and cables.

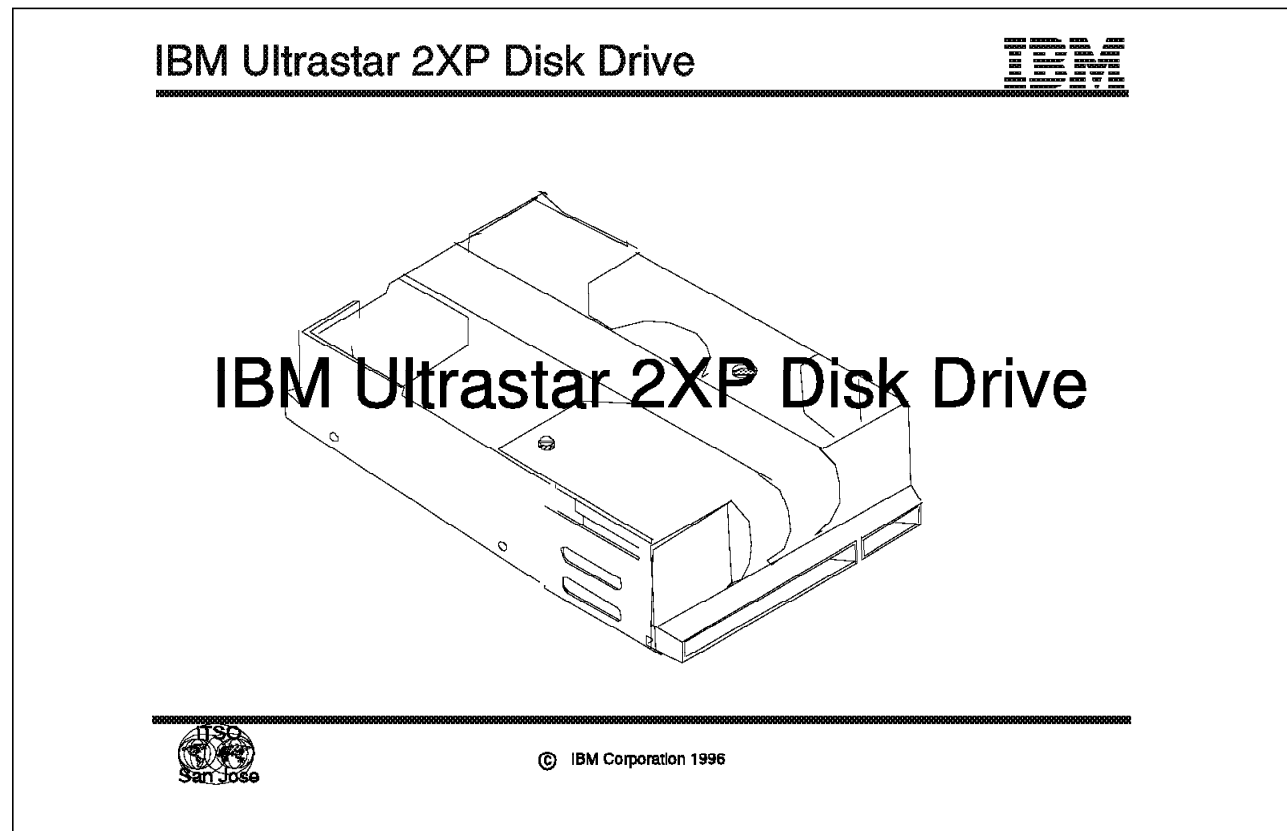
Third generation RAID-5

The RAMAC 3 Array drawer is the third generation of IBM enhanced RAID-5 high-end DASD storage products. It builds on the knowledge and experience gained from the implementation of its predecessor products, RAMAC and RAMAC 2. The RAMAC 3 Array drawer capitalizes on the innovations and progress made in the field of disk drive technology and architecture.

Fault-tolerant design

High availability is delivered through a fault-tolerant hardware design that guards against any single component failure jeopardizing data integrity, availability, and access.

Chapter 4. IBM Ultrastar 2XP Disk Drive

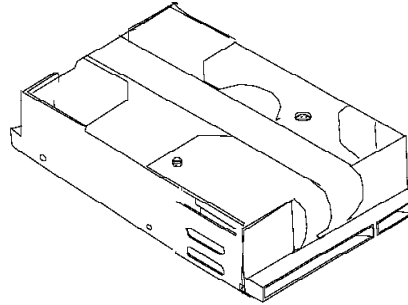


The driving force behind many recent IBM storage subsystem enhancements is the remarkable progress made in the field of disk drive technology and architecture. Improvements in disk drive technology and architecture inevitably lead to increased capacity and improved reliability that simply cannot be ignored in the quest to satisfy marketplace demand for cost-effectiveness, continuous availability, and growing functionality. The RAMAC Array Family has spanned a three-year period characterized by phenomenal changes in disk drive technology and architecture. The RAMAC Array Family has adapted to and exploited changes in technology and architecture such as 3.25 in. form factor disks, zoned bit recording, embedded servo, MR head, PRML channel, and No-ID sector format. This chapter describes the new Ultrastar 2XP disk drive functions and features that are exploited by RAMAC 3 Array Storage.

IBM Ultrastar 2XP Disk Drive ...



- 9.22 GB
 - 9 disks and 18 heads
 - 6160 tracks per inch (TPI)
 - 828.7 Mbit/sq in.
- 7202.7 rpm
- No-ID sector format
- Zoned bit recording
- MR heads
- 1 MB disk cache
- Predictive Failure Analysis
 - S.M.A.R.T compliant



© IBM Corporation 1996

IBM Ultrastar 2XP Disk Drive

The Ultrastar 2XP disk drive is IBM's newest high-capacity, high-performance disk drive. It uses the latest technology to achieve a high capacity in a small form factor. There are various models of the Ultrastar disk. The model used in RAMAC 3 has been optimized to operate in the high-availability environment of the System/390 customer.

9.22 GB

The Ultrastar 2XP disk drive has a capacity of 9.22 GB. This capacity is achieved by using 9 disk platters with 18 heads. A linear density of between 116,585 and 134,544 bits per inch and 6160 tracks per inch gives an areal density of up to 829 Mbit/sq in. This density is more than 50% greater than the areal density of the disk used in RAMAC 2, which has an areal density of 544 Mbit/sq in.

7200 rpm

A rotation speed of 7200 rpm enables fast access to data on a track, reducing overall response times.

No-ID sector format

No-ID sector format is a new innovation recently introduced on IBM disks. It enables higher capacities on disks and faster handling of spare blocks (see foil “Ultrastar 2XP No-ID Sector Format” on page 62 for more details).

Zoned bit recording

Zoned bit recoding takes advantage of the fact that the track lengths of a disk differ significantly between the inner tracks and the outer tracks. Rather than use a constant number of bytes per track (as is used on the 3390, for example) where the linear density on the inner tracks is higher than on the outer tracks, zoned-bit recording keeps the density the same and varies the amount of data on each track. On an Ultrastar 2XP disk drive, the number of data sectors per track varies between 120 sectors on the inner tracks and 184 sectors on the outer tracks. The main benefit of using zoned-bit recording is increased capacity. An interesting side benefit is that the outer tracks have a higher data rate than the inner tracks. On the Ultrastar 2XP disk drive, the disk is divided into eight areas or zones. However, almost 60% of the data resides in the first two zones on the outer part of the disk that gives the highest data rate.

MR heads

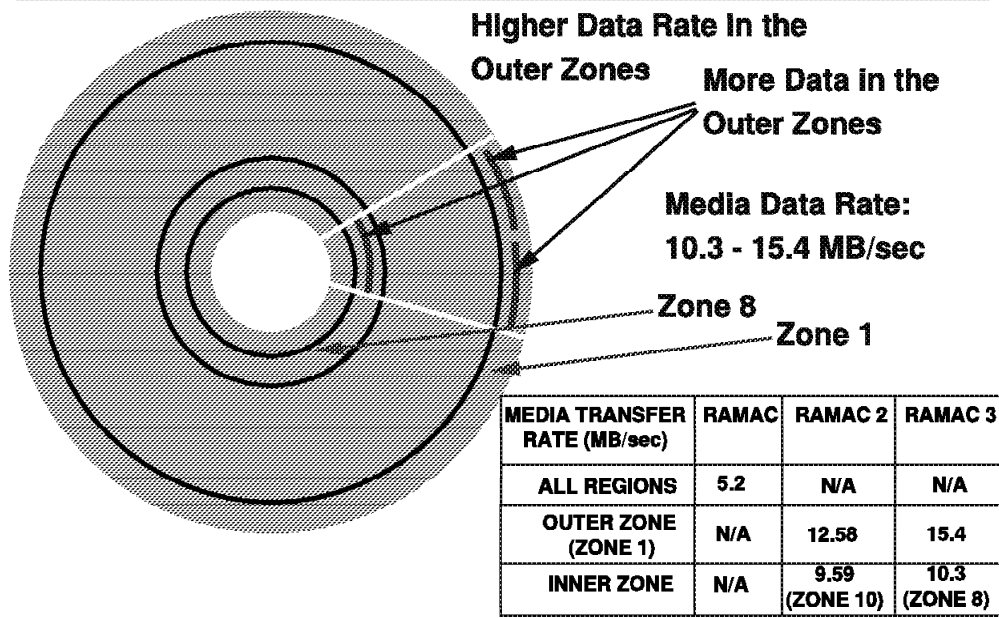
The Ultrastar 2XP disk drive disks use IBM’s magneto-resistive (MR) head technology, which enables IBM to develop disks with high areal density. Thus IBM can produce high capacity disks with a minimum number of disk surfaces and heads, keeping costs down.

Predictive Failure Analysis - S.M.A.R.T compliant

IBM Predictive Failure Analysis (PFA) complies with the Self-Monitoring, Analysis and Reporting Technology (S.M.A.R.T.) industry standards.

PFA routines, invoked when the disk is idle, measure the characteristics of each head, saving the data to a log in a reserved area of the disk. The log data can be retrieved and analyzed. Should any characteristics exceed a specified threshold, the RAMAC drawer is notified and a service information message (SIM) is issued. The RAID-5 function of RAMAC protects customer data, and a failing disk can be replaced and rebuilt in place using dynamic device reconstruction (DDR). PFA is detailed in “IBM Ultrastar 2XP Disk Drive Reliability” on page 70

RAMAC 3 - Zoned Bit Recording



© IBM Corporation 1996

RAMAC 3 - Zoned Bit Recording

The Ultrastar 2XP disks used in the RAMAC 3 drawer are partitioned into recording bands or zones (also sometimes referred to as *notches*). There are eight zones on each disk; zone 1 is located nearest the outer diameter of the disk, and zone 8 is located nearest the inner diameter of the disk. With zoned bit recording, the inner and outer diameter regions of the disk have the same areal density (the same number of megabits per square inch).

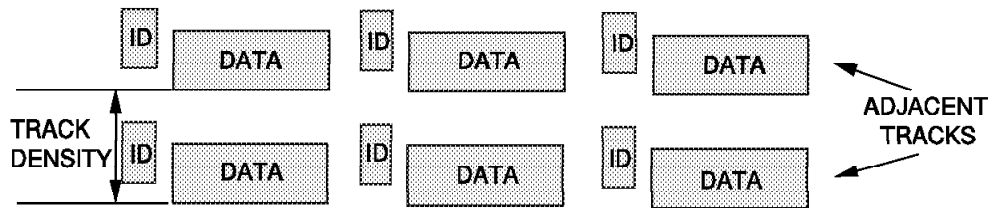
The RAMAC 2 drawer was the first member of the RAMAC Array Family to use zoned bit recording disks. (Previously, the areal density was not the same in the inner and outer diameter regions of the disk because the disk drive channel technology could not support the faster radial speeds of the outer zones.) With zoned bit recording, more data can be stored in an outer zone than an inner zone. Thus the capacity of the disk is increased because there is no longer a need to "space out" the data in the outer zones because of radial speed. The new generation of partial response maximum likelihood (PRML) disk drive channels supports the higher radial speeds of zoned bit recording disks.

Because the radial speed of the outer zones is about twice as high as the inner zones, the transfer rate is higher in the outer zones. The variation in media transfer rate is shown on the foil for the innermost and outermost zones for RAMAC 2 and RAMAC 3.

The Ultrastar 2XP disk used in the RAMAC 3 drawer has a block format length of 688 bytes (this is sometimes referred to as the *SCSI sector* or *block length*).

There are 13,408,182 of these blocks on the drive, giving a capacity of 9,224,829,216.

ID Sector Format



- ID field needed to identify data
 - Logical sector number
 - Sector status
 - ▶ *good*
 - ▶ *bad - pointer to alternate sector*
 - CRC
- Presence of ID field reduces track density with MR head
 - MR head separates read and write elements
 - ID field offset from data to enable read head alignment
- Presence of ID field reduces track capacity



© IBM Corporation 1996

ID Sector Format

The Ultrastar 2XP disk drive uses a disk architecture different from typical disk drives in the marketplace today. The Ultrastar 2XP disk used in the RAMAC 3 drawer implements the No-ID sector format architecture. The No-ID sector format increases the capacity of the disk by removing the ID field or header information from the disk surface. To better understand this disk architecture innovation, we will first look at the predecessor technique, ID sector format, that is used in RAMAC 2 storage subsystems.

ID field needed to identify data

Disk drives implementing the ID sector format architecture require that an ID field be written to the disk ahead of the data sector. This ID field or header information identifies the data sector through its logical sector number. When the host or control unit reads or writes data, it sends a list of logical block numbers to the disk drive, and the disk drive converts these numbers into zone, cylinder, head, and sector values (ZCHS). After the drive seeks to the desired zone, cylinder, and head, the ID fields on the disk surface are read until a match is found with the ZCHS value. The data field that follows (this ID field) can then be read or written. In addition to identifying the data sector, the ID field also contains information about the status of the sector. If the sector is a bad sector, the data is reassigned to another part of the disk so that successful read and write processing can progress for future accesses to that data. The ID field is therefore an essential component for normal data access as well as for defect management such as defect sector reassignment.

Presence of ID field reduces track density with MR head

The foil shows the location of the ID field in relation to the data field, but it also highlights an important aspect of how the ID field is aligned on the track in relation to the data field. Notice that the ID field is offset from the data field. The apparent misalignment of the fields is deliberate, the result of separating the read and write elements on the MR head. (See next foil.)

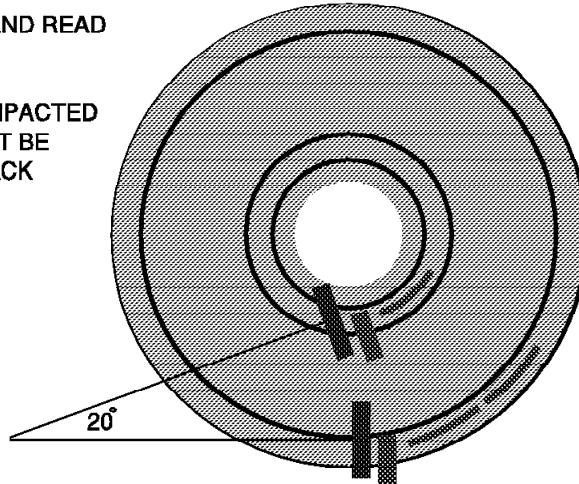
Presence of ID field reduces track capacity

The ID field occupies space on the disk surface—every data field must be preceded by an ID field. The ID field uses approximately 5%-15% of the disk capacity, but this space is needed because of the important role that the ID field plays in data access and defect management. It is clear, however, that any advance in disk drive architecture that enables us to reclaim this capacity would be welcome and beneficial. The next foil describes the effect of MR head technology on the ID sector format.

MR Head Geometry and ID Sector Format



- ROTARY ACTUATOR MOTION ACROSS TRACK
SKEWS HEAD AND TRACK ALIGNMENT
- CENTERLINE OF WRITE HEAD AND READ
HEAD OFFSET
- OPERATIONS AND CAPACITY IMPACTED
BECAUSE BOTH HEADS CANNOT BE
ALIGNED WITH CENTER OF TRACK



© IBM Corporation 1996

MR Head Geometry and ID Sector Format

A major development in the field of disk drive capacity is the MR head. The MR head enabled areal densities to grow to 268 Mbit/sq in. on the 3390-9; 259 Mbit/sq in. on the 3.5 inch form factor disk used in RAMAC; 544 Mbit/sq in. for RAMAC 2; and 829 Mbit/sq in. on RAMAC 3. These developments in areal density facilitated the explosive growth in device capacities from 2.89 GB (3390-3) to 9.22 GB (RAMAC 3) while reducing form factor and environmental requirements. Innovations in thin-film disk technology, PRML channel technology, and disk drive architecture accompanied the MR head developments.

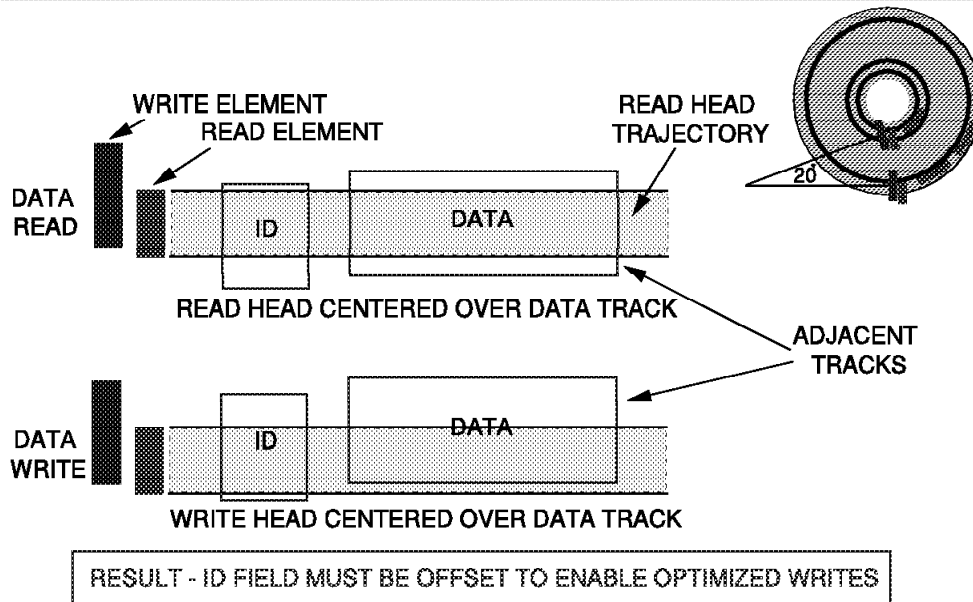
Optimal writing requires a head with only a few thick wire coils, a high current, and slow rotation of the disk. Optimal reading, however, requires a head with many thin wire coils and a high rotation speed to produce a high induction signal. Because reading and writing have contrary physical requirements, the conventional inductive read-write head is a compromise between these requirements. The MR head derived much of its advantages through separation of the read and write elements based on the above observations. The gains made by enabling dense packaging of the data meant that MR head technology and associated geometry should be supported by complementary technologies and architectures.

The foil shows the relationship between the MR head and the motion of the rotary actuator as it moves across the surface of the disk. The rotary actuator has a swing angle of approximately 20 degrees and results in the heads being skewed in relation to the data tracks. Because optimum performance is achieved by centering the appropriate head over the data track, this geometric

effect impacts capacity and read-write operations—both heads cannot be aligned with the center of the data track. To compensate for the offset of the read and write head, the ID field used in ID sector format is also offset to enable the data sector to be read and identified before a write operation. The next foil describes this effect in more detail.

MR Head Effect on ID Sector Format

IBM



© IBM Corporation 1996

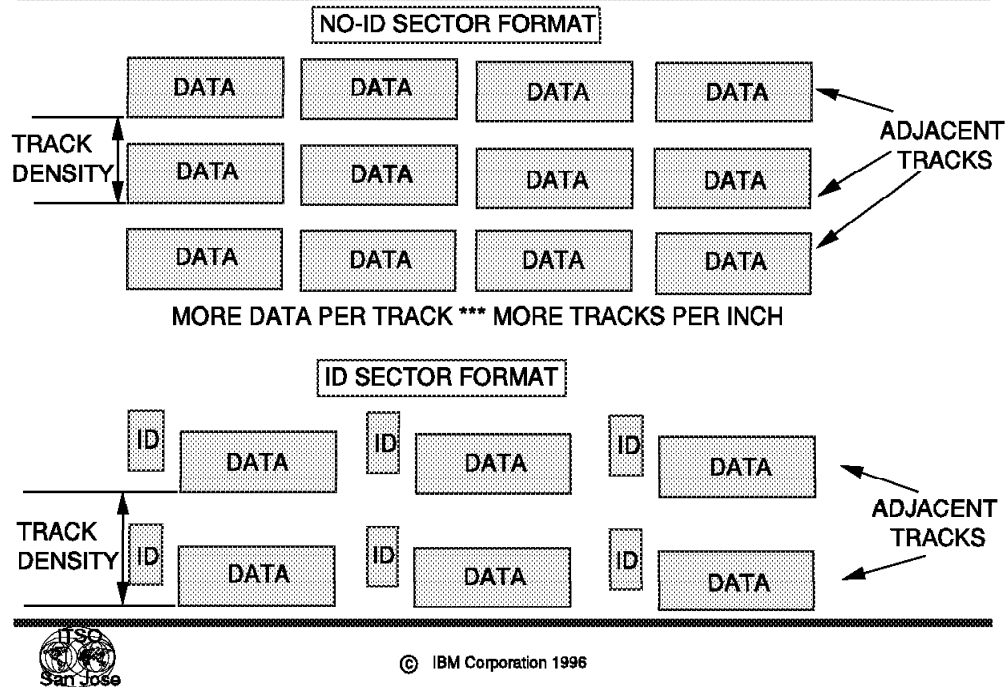
MR Head Effect on ID Sector Format

The ID field must be read before the data field can be read or written. This mandates that the read head be positioned over the ID field before read and write operations. For a read operation, the read head should be positioned over the center of the data track. For a write operation, the write head should be positioned over the center of the data track. In both cases, the read head must first be able to read the ID field.

The problem and the solution are highlighted on the foil in the Data Write example. For a data write operation, the write head must be aligned with the center of the data track, but the alignment offsets the read head to a position "offtrack"—the position of the read head makes it impossible to read the data track. For a write operation, however, the read head only has to be able to read the ID field, so the ID field is offset to enable the read head to read the ID and the write head to write the data while the MR head is in the same position. Thus "read ID before write data" can take place and is in essence the ID sector format architecture accommodating the rotary actuator and MR head technologies to bring high-performance, high-capacity disk drives to the marketplace. The gains in areal density through MR head exploitation make the ID sector format an architecture worth adopting to ensure highly reliable data accesses. However, track density is limited because we need a wider write head.

The next challenge is to implement an architecture that removes the ID field and thus the need to "offset" it.

No-ID Sector and ID Sector Format



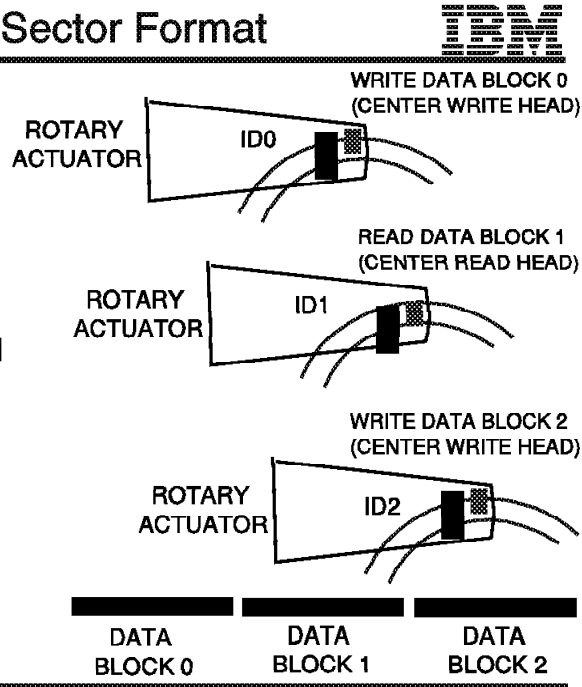
No-ID Sector and ID Sector Format

The foil compares the No-ID sector format and the ID sector format. By removing the ID fields from the disk surface, more data can be stored on each track. This increase in track capacity accounts for approximately 5%-15% of disk capacity. In addition the track density is improved because tracks can be written closer together because the "offset" ID field has been removed and the write head is 15% narrower. This reclamation of capacity through increased track density accounts for approximately another 15% of disk capacity.

The foil underscores the significant impact of No-ID sector format, which provides more data per track and more tracks per inch. But where did the ID information go? Having established the necessity of the ID field for data sector identification and defect management, can it simply be discarded so that disk space can be reclaimed? The answer is no. The next foil describes the key aspects of the No-ID architecture as implemented in the Ultrastar 2XP disks in the RAMAC 3 drawer.

Ultrastar 2XP - No-ID Sector Format

- ID information computed from disk drive solid-state memory
 - In compressed format
 - *Rapid lookup*
 - *Small size*
- Servo and data control systems perform high-speed address generation
 - LBN to PBN translation
 - Defect management
 - Spare block assignment
- Eliminates read ID before read and write operations
 - Sector location computed in advance



© IBM Corporation 1996

Ultrastar 2XP No-ID Sector Format

Typical disk drives today have ID fields associated with each data sector. The ID fields are used to identify data sectors and for defect management. In the No-ID sector format, these functions are now performed electronically. The actual ID information is no longer stored on the disk.

ID information computed from disk drive solid state memory

Header information formerly stored in the ID field on the disk is now calculated from data stored in solid state memory on the disk drive. The disk drive memory contains a “defect map” that is used to locate the exact location of the requested data sector—even if the sector has been reassigned to a different physical location. Storing the ID information in a semiconductor storage medium provides fast access to the information in advance of positioning the read-write heads. The data sector identification information is stored in compressed format in the drive memory, thus enabling rapid lookup and efficient memory space utilization.

Servo and data control systems perform high-speed address translation

The disk drive must convert the logical block numbers (LBNs) supplied by the host through the 3990-6 or 9390 and RAMAC 3 drawer to physical block numbers (PBNs). The PBN takes account of any sector reassignments that have resulted from bad blocks on the disks. Because the defect map is already computed, the PBN can be converted to the exact zone, cylinder, head, and sector values that locate the requested data field. The LBN-to-PBN conversion is performed by the Ultrastar 2XP disk drive servo and data control system.

Eliminates read ID before read and write operations

By holding the ID in the disk drive memory, the read of the ID field that is normally required before a write operation is eliminated. Instead the disk drive electronics calculate the exact location of the requested data sector by using the ID information stored in the disk drive memory. Thus the location of the data is known in advance of rotary arm movement and the need to read and compare the ID field on the disk surface is avoided.

More information about No-ID formatting can be found on the IBM Almaden Research web site.

<http://www.almaden.ibm.com/sst/diskarch/index.htm>

For more detailed information

<http://www.storage.ibm.com/storage/oem/tech/noid.htm> — Introducing the No-ID Sector Format.

Advantages of No-ID Sector Format



- Increased disk capacity
 - Increased track capacity
 - *Up to 15%*
 - Increased track density
 - *Up to 15% of disk capacity reclaimed*
- Improved performance
 - Improved access times
 - *Avoid "read ID field" overhead*
 - Increased throughput
 - *More data on track*
- Improved defect management
- Improved power management
 - Avoid powering read electronics during sector search
- Enhanced reliability
 - Header data stored in memory
 - Avoid CRC or ECC processing



© IBM Corporation 1996

Advantages of Ultrastar 2XP No-ID Sector Format

The advantages of the Ultrastar 2XP disk drive No-ID sector format architecture contribute to RAMAC 3 Array Storage's cost-effective, high-capacity performance, availability, and reliability design objectives. The foil summarizes the benefits of IBM's unique No-ID sector architecture.

Increased disk capacity

The removal of the ID field enables IBM to increase the number of sectors on a track. Removal of the ID field does not influence linear density (number of bits per inch along a track) or areal density (number of megabits per square inch on the disk), but it reclaims valuable disk capacity that previously was unusable because of architectural requirements in securing highly reliable data accesses.

With an MR head, the read and write elements are separate, and head skew occurs as the rotary actuator moves across the disk surface. A normal write must read the ID field immediately preceding the sector to be written, and a compromise must be reached between the position of the read element and the write element that allows a read immediately followed by the write. In a disk without No-ID recording format, the read and write heads are offset to compensate, thus limiting the track density that can be achieved. With No-ID recording, only an alignment for writing takes place as the ID is now in the disk drive memory. Thus the offset in read and write elements is less important and tracks can be closer together.

No-ID recording increases track capacity by up to 15% and allows tracks to be closer together by 15%, giving an overall 30% improvement in disk capacity.

Improved performance

The performance of the Ultrastar 2XP disk drive is improved through use of the No-ID sector format architecture because data fields can be read or written faster as a result of computing the correct sector location in advance of head positioning. Previously, with the ID sector format, access times included the time to read and compare the ID field for both read and write operations. Remember that the disk drive must validate the correct data sector before a read and write operation. Access times are therefore improved when the No-ID sector format architecture enables "data sector validation" electronically without the need to read preformatted fields on the disk surface.

The No-ID sector format also increases the throughput of the disk drive because more real data can be transferred per read and write operation. Read and write operations are concentrated on accessing real data as opposed to nondata header fields.

Improved defect management

Before the implementation of the Ultrastar 2XP No-ID sector architecture, the ID field stored on the disk surface contained the details of sector status and the location of any reassigned data blocks. For this information to be retrieved and used to read and write data, the ID field had to be read first. By storing sector defect maps in disk drive solid state memory, the defect management process is executed at electronic speeds without requiring physical disk access.

Improved power management

The Ultrastar 2XP disk drives in the RAMAC 3 drawer use a power management system that detects component-idle periods. During periods when certain components are not used, the Ultrastar 2XP disk drive intelligently avoids supplying power to those components. For example, during data reads, power is required to sustain the read operation, but when the data is not being read, the read mechanisms and electronics are not used (thus saving power).

Consider the No-ID sector format architecture and the synergy with the Ultrastar 2XP power management system. The No-ID sector format obviates the need to perform a read of the ID field before read and write operations. Thus the read electronics do not need power during the sector search that would normally have preceded these operations.

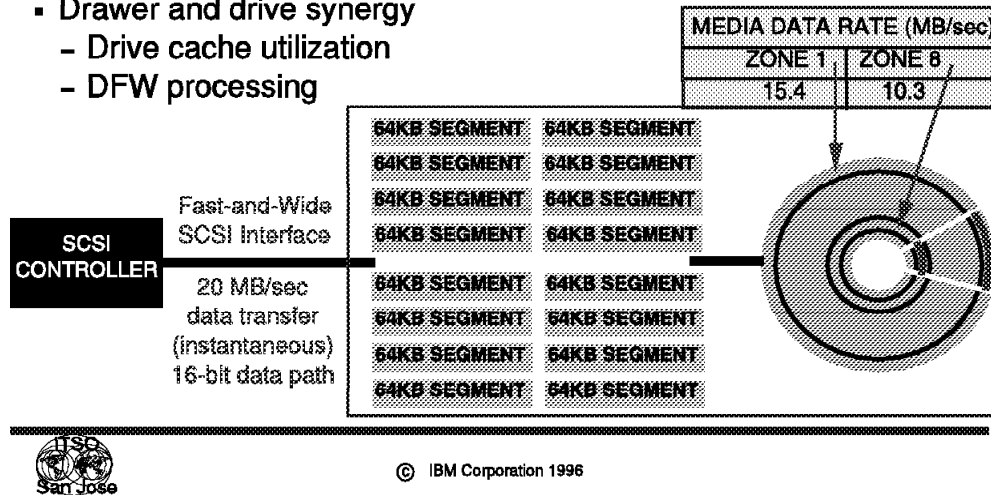
Enhanced Reliability

Random access memory (RAM) has a lower error rate than magnetic disk. Therefore storing the ID field in solid state memory on the disk drive reduces the vulnerability to errors in the ID field. Cyclic redundancy checks (CRCs) or error checking and correction (ECC) techniques are used effectively by traditional disk drives to protect the ID field contents stored on the disk surface. No-ID sector format avoids the recovery processes associated with CRC or ECC and further improves disk drive reliability.

RAMAC 3 Disk Drive Cache



- 1 MB disk drive cache
- Cache segment size doubled
 - Segments provide data integrity checking (CRC)
- Drawer battery provides nonvolatility
- Drawer and drive synergy
 - Drive cache utilization
 - DFW processing



© IBM Corporation 1996

RAMAC 3 Disk Drive Cache

1 MB disk drive cache

Each Ultrastar 2XP disk drive has a 1 MB cache that is used for read and write data transfer operations with the RAMAC 3 drawer. The disk drive cache is partitioned into sixteen 64 KB segments—the disk drives used by RAMAC and RAMAC 2, although different, both use a 512 KB disk drive cache that has sixteen 32 KB segments.

Cache segment size doubled

By doubling the disk drive cache segment size, more data can be written from the RAMAC 3 drawer during destage operations. The 64 KB segments include built-in data integrity checks by using CRC codes. Performing CRC processing on disk drive cache segments enables the disk drive controller to detect and correct certain types of bit pattern inconsistencies during data transfer to and from disk drive cache.

Drawer battery provides nonvolatility

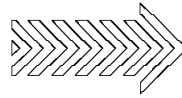
The drawer battery protects the disk drive cache contents in the event of drawer power loss. Thus the disk drive cache can be considered nonvolatile storage. Each RAMAC 3 drawer therefore contains 68 MB of nonvolatile storage optimized for enhanced RAID-5 processing and intelligent destage management. The disk drive cache is also used for data prestaging.

Drawer and drive synergy

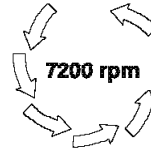
The RAMAC 3 drawer implements sophisticated RAID-5 processing operations, among which are the unique destage techniques described in Appendix A. The RAMAC 3 drawer writes data asynchronously to the disk drives and ensures that lower interface utilization is optimized by taking advantage of the disk drive cache for "grouped destage" events. Two significant functions enable drawer-to-drive data transfer efficiency. The first is the ability of the RAMAC 3 drawer to determine the utilization of the disk drive cache segments before grouping updates for destage. The RAMAC 3 drawer and the Ultrastar disk drive perform a "look-ahead" function that informs the drawer of disk drive cache segment occupancy and thus enables the RAMAC 3 drawer to schedule destage activity to the disks. The second function is the ability of the RAMAC 3 drawer to exploit the nonvolatile characteristics of the disk drive cache. The RAMAC 3 drawer can initiate destage operations to the drive cache at optimal SCSI data transfer speeds and free up the drawer-drive interface faster for other I/O destage activities. The disk drive controller performs the write from disk cache to physical disk asynchronously with RAMAC 3 drawer destage operations, providing further overlap with other drawer data transfer activity on the lower interface.

IBM Ultrastar 2XP Disk Drive - Performance

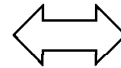
- 20 MB/s Fast-and-Wide SCSI
- Media data rate 10.3-15.4 MB/s
- Average seek (read) 8.5 ms
- Track to track seek 0.5 ms
- Latency 4.17 ms



10.3 - 15.4 MB/s



7200 rpm



8.5 ms



© IBM Corporation 1996

IBM Ultrastar 2XP Disk Drive - Performance

This foil summarizes the performance characteristics of the Ultrastar 2XP disk drive.

20 MB/s Fast-and-Wide SCSI

The fast 20 MB/s interface provides for high-performance disk drive cache to drawer cache transfers. Many of the RAMAC 3 Array Storage processing tasks are asynchronous, resulting in a “decoupling” of data transfer throughout the subsystem hierarchy. The fast 20 MB/s SCSI enables optimized data transfer for lower interface activity.

Media data rate 10.3-15.4 MB/s

The Ultrastar 2XP has a fast transfer speed, which is important where direct disk access is required (compare this with 4.2 MB/s on the 3390). When data requests cannot be satisfied from either subsystem or drawer cache, physical disk access is required. The Ultrastar 2XP supports media data rate transfers of between 10.3 and 15.4 MB/s depending on which zone of the disk the data is located. The high-speed media data rate of the Ultrastar 2XP ensures the highest levels of disk drive performance when application accesses result in data transfer from the physical media.

Average seek (read) 8.5 ms

Low seek times enable fast access to the large amount of data on the 9 GB drives when requests are not satisfied by subsystem or drawer cache.

Track to track seek 0.5 ms

Most seek activity is to adjacent tracks—the Ultrastar 2XP has fast track to track seek times.

Latency 4.17 ms

The rotation speed of 7200 rpm gives a latency of only 4.17 ms, which together with the fast seek times give the drive excellent service times when physical disk access is unavoidable because of subsystem and drawer cache miss.

IBM Ultrastar 2XP Disk Drive - Reliability



- PFA
 - Analyses drive parameters
 - Predicts imminent drive failures
 - Idle-time function
 - ▶ *7 measurements on each head*
 - ▶ *4-hr cycle*
 - ▶ *Duration is 80 ms*
- Channel calibration
 - Calibrates read/write circuits
 - Idle-time function
 - ▶ *Executed if drive is idle for 5 s*
 - ▶ *4-hr cycle*
 - ▶ *Duration is 20 ms*



© IBM Corporation 1996

IBM Ultrastar 2XP Disk Drive Reliability

The Ultrastar 2XP performs many tests on disk drive components to ensure that they are not only functioning but functioning optimally. The tests are designed to measure disk drive parameters and predict whether component failures are imminent.

PFA

Predictive failure analysis (PFA) is an idle-time function that consists of seven measurements taken for each head on the disk. Idle-time functions are the least intrusive way of testing for potential component malfunctions because they activate during times when the disk drive components are idle—production I/O processing preempts idle-time function execution. The seven PFA measurements are taken for each head over a 4-hr period, and each takes approximately 80 ms. The disk drive spreads the PFA idle-time measurements evenly throughout the 4-hr period. For the Ultrastar 2XP used in the RAMAC 3 drawer, one PFA measurement is performed about every 2 min, thus ensuring the highest levels of disk drive head error detection and warning.

Channel Calibration

The Ultrastar 2XP periodically calibrates the channel (the disk drive channel) to ensure that the read and write circuits function optimally, thus reducing the likelihood of soft errors. Like the PFA function, channel calibration is performed every 4 hr and typically completes in 20 ms. The calibration exercise starts only if the disk drive has been idle for 5 s, again ensuring that there is no interference with production I/O processing.

IBM Ultrastar 2XP Disk Drive - Reliability...



- Log recording
 - Saves periodic data on reserved area of disk
 - Idle-time function
 - ▶ Executed if drive is idle for 5 s
 - ▶ Logs saved every 26 to 35 min
 - ▶ Duration is less than 30 milliseconds
 - Data used for failure analysis
- Disk sweep
 - Idle-time function
 - ▶ Executed if drive is idle for 40 s
 - Heads moved after 9 min in same location



© IBM Corporation 1996

Ultrastar 2XP Reliability...

Log recording

The Ultrastar 2XP disk drive periodically saves data in logs located on reserved areas of the disks. These reserved areas of the disk do not encroach on the useable capacity available for customers' application data. Log recording is an idle-time function that is invoked only if the disk drive has been idle for 5 s or more between SCSI command executions. Log recording takes approximately 30 ms, and the logs are saved about every half-hour. The data recorded in the logs is used for failure analysis.

Disk sweep

If the Ultrastar 2XP disk drive has not processed a SCSI command for at least 40 s, the drive executes the disk sweep idle-time function. Disk sweep exercises the heads by moving them to another area of the disk, and it initiates a second move of the heads if they fly in the same spot for 9 min. The disk sweep idle-time function ensures that low-use drive components do not become vulnerable to errors through inactivity.

IBM Ultrastar 2XP Disk Drive - Summary



HIGH PERFORMANCE

- Rotational speed: 7202 rpm
- Data rate: 10.3-15.4 MB/s
- Disk drive cache
- Fast-and-wide SCSI
- No-ID sector format
 - Access time
 - Throughput

HIGH RELIABILITY

- PFA
- No-ID sector format
 - Defect management

HIGH CAPACITY

- Areal density
 - Third generation
 - *MR head*
 - *PRML channel*
- Zoned bit recording
- Embedded sector servo
- No-ID sector format
 - Track capacity
 - Track density

OPERATING COSTS

- Power management



© IBM Corporation 1996

IBM Ultrastar 2XP Disk Drive - Summary

The foil summarizes the benefits of the Ultrastar 2XP disk drive in the categories of high performance, high capacity, high reliability, and operating cost.

High performance

RAMAC 3 Array Storage subsystems implement a multilevel cache hierarchy that exploits cache memory in the storage control and drawer. The Ultrastar 2XP disk drive expands the cache hierarchy by providing 1 MB of cache with each disk drive (4 MB of cache for each RAMAC 3 drawer) to improve I/O service times and asynchronous destaging. When data access cannot be satisfied by using storage control, drawer, or disk cache, the rotational speed, data rate, and high-speed SCSI interface of the Ultrastar 2XP disk drive ensure optimum physical media performance. The new No-ID sector format architecture used in the Ultrastar 2XP disk drive enables greater throughput and faster access to data by avoiding ID field management and processing that require extra physical media accesses.

High capacity

The 9.22 GB Ultrastar 2XP disk drive incorporates many functions that deliver high-capacity data storage in a 3.25 in. form-factor disk drive. The areal density of the Ultrastar 2XP disk drive is more than 50% greater than that of the disk drive used in the RAMAC 2 drawer. The areal density improvements are complemented by the latest generation of MR head and PRML technologies. Add to this the unique combination of zoned bit recording and embedded servo (first pioneered with RAMAC 2) and No-ID sector format, and it comes as no

surprise that the Ultrastar 2XP is the industry-leading physical media for storage subsystems.

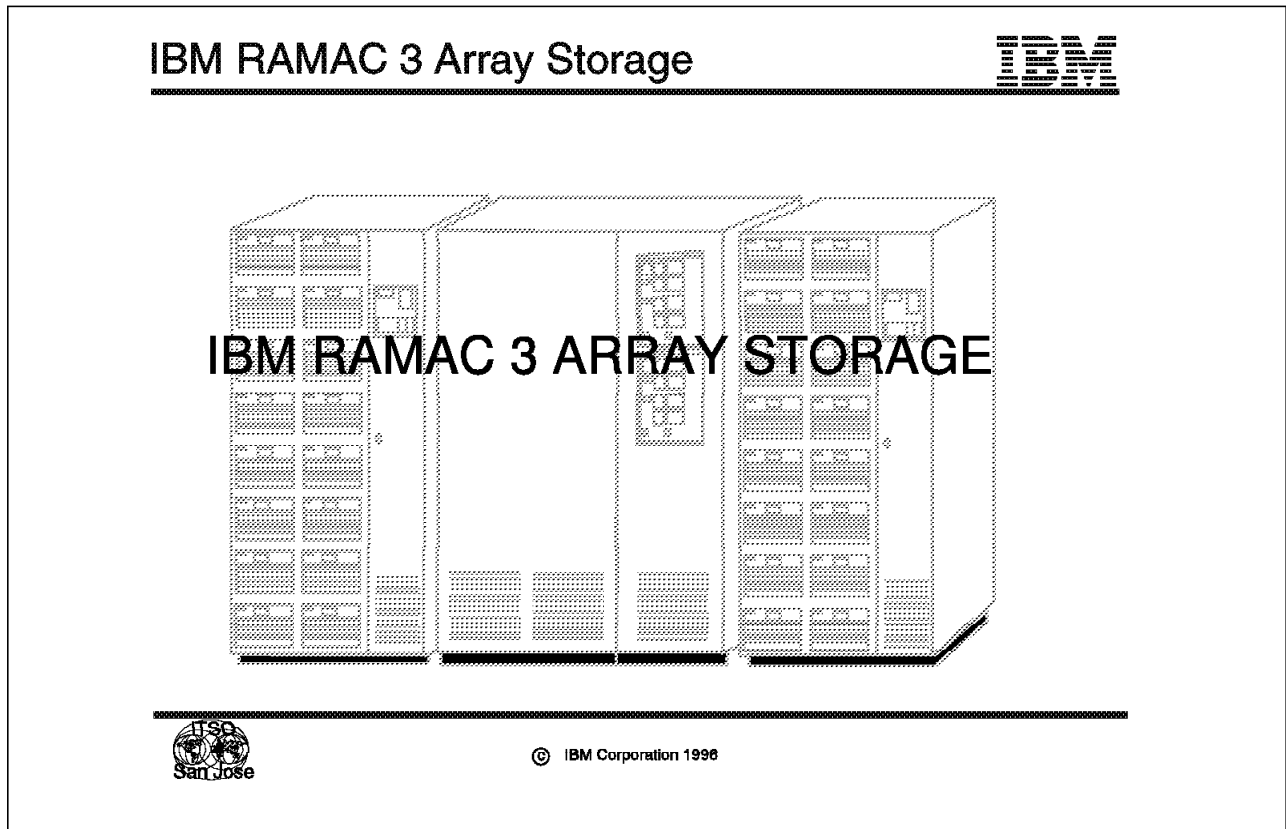
High reliability

The likelihood of disk drive failure is minimized with the Ultrastar 2XP disk drive. The disk drive implements PFA functions that enable early detection of component malfunction in advance of application I/O. Idle-time functions ensure that low-usage drive mechanics and electronics are exercised and monitored. No-ID sector format enables faster and more reliable defect management procedures by avoiding disk drive access if sector reassignment is necessary.

Operating cost

The Ultrastar 2XP disk drive reduces the operating cost of RAID-5 data storage by using intelligent electronics that dynamically adjust power consumption based on component requirements. Advanced power management is even more significant when we realize that the number of Ultrastar 2XP disk drives in a fully configured RAMAC 3 Array Storage subsystem is 128.

Chapter 5. IBM RAMAC 3 Array Storage



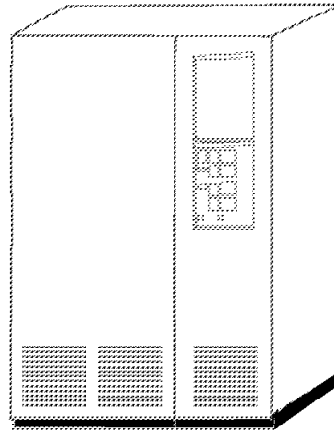
IBM RAMAC 3 Array Storage

This section covers the new RAMAC 3 Array Storage subsystem, which consists of two components, the RAMAC 3 Storage Control and one or two RAMAC 3 Storage Frames. The RAMAC 3 Storage Control has a device type of 9390. There are two models, the 9390-1 and the 9390-2.

9390 Model 1 Storage Control



- Single storage control
- Supports 128 devices
- Cache
 - From 256 MB to 4 GB
- Nonvolatile storage
 - From 32 MB to 128 MB
- Channels
 - ESCON only
 - 4 to 16 ESCON
- Basic mode default



© IBM Corporation 1996

9390-001 Storage Control

The 9390-001 is a new IBM storage control that attaches to a single RAMAC 3 Storage Frame. It is the same height and color as the RAMAC 3 Storage Frame.

Single Storage Control

The 9390-001 contains a single storage control providing cache, nonvolatile storage, ESCON channels, and four storage paths. It implements a fault-tolerant design providing all of the extended functions of the 3990-6. The 9390-001 contains a maintenance processor in each cluster (similar to the support facility (SP) in the 3990), and shared control arrays (SCAs) used for storing and managing extended function status.

Supports 128 devices

One of the major enhancements over currently installed 3990-6s is the support for 128 RAMAC 3 devices. This support enables a storage control to address 128 logical 3390-3 volumes, doubling the current limit of 64 volumes.

Cache

Cache sizes for the 9390-001 start at 256 MB and go up to 4 GB. The sizes are:

- 256 MB
- 512 MB
- 768 MB
- 1024 MB
- 1536 MB

- 2048 MB
- 3072 MB
- 4096 MB

The workload characteristics, I/O rate, and required levels of performance determine the actual size. IBM has modeling tools to assist you in determining the most suitable size for your workload.

Nonvolatile storage

Three nonvolatile storage sizes are available on the 9390-001: 32 MB, 64 MB, and 128 MB. When choosing an NVS size, consider the read-to-write ratio of the workload.

Channels

The 9390-001 supports ESCON channels only. If you want to support RAMAC 3 Array drawers from a parallel channel host, you can attach a RAMAC 3 Storage Frame to a 3990-6, which supports parallel channels. There are several channel combinations, which are made up of either two-port ESCON cards or four-port ESCON cards. In both cases each ESCON card supports 32 logical paths.

Valid combinations are:

- 4 ESCON - (2x2-port cards) - 64 logical paths
- 8 ESCON - (4x2-port cards) - 128 logical paths
- 8 ESCON - (2x4-port cards) - 64 logical paths
- 12 ESCON - (2x2-port cards and 2x4-port cards) - 128 logical paths
- 16 ESCON - (4x4-port cards) - 128 logical paths

If you are using ESCON Directors, you probably need only four ESCON ports. If you are using PPRC or plan to need more than the minimum four ports, 4-port cards provide expansion up to the maximum 16 ports.

Basic mode default

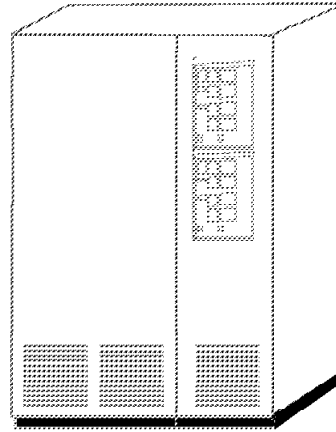
The 9390 defaults to basic mode as opposed to enhanced mode. From an operational point of view there is little difference between the two modes. On the 3990-6, enhanced mode supports rotational position sensing (RPS) miss avoidance on parallel channels—this does not apply to a 9390, which is ESCON only. The second function that enhanced mode supports is control unit initiated reconfiguration (CUIR). The service representative uses CUIR to take host channels offline and online through the 3990 or 9390 panel as part of a service procedure. PTFs may be required to your MVS system to support CUIR. A 9390 in enhanced mode responds to a host request for device characteristics, saying that it is a 3990-6. In basic mode a storage control responds that it is a 3990-3.

When you use dual copy on a 9390 with 128 addresses, unless the 9390 is in enhanced mode, some of the dual copy messages are garbled. So, if you do not plan to use dual copy, you can install your 9390 in basic mode.

9390 Model 2 Storage Control



- Dual storage control
- Supports all extended functions
- Supports 256 devices
- Field upgrade from 9390-1 to 9390-2
- Cache
 - 512 MB to 8 GB
- Nonvolatile storage
 - 64 MB to 256 MB
- Channels
 - 8 to 32 ESCON
- Common power supplies



© IBM Corporation 1996

9390-002 Storage Control

The 9390-002, together with two Storage Frames of RAMAC 3, provides a dramatic improvement in gigabytes per square foot, when compared with the 3990 and RAMAC 2.

Dual storage control

The 9390-002 is a dual storage control. It can be considered as two 9390-001s in a single box. A 9390-001 can be upgraded to a 9390-002 by the addition of the second storage control. This upgrade can be performed concurrently with production use of the currently installed storage control.

Each storage control attaches to a RAMAC 3 Storage Frame through new director-to-device connection (DDC) cables. The RAMAC 3 Storage Frames can be located anywhere within a distance of 200 ft. from the 9390.

Supports all extended functions

All functions available on the 3990-6 are available on both storage controls in the 9390-002 box. Each storage control in a 9390-002 should be considered as completely independent. Therefore when configuring a 9390-002, configure it as if it were two 9390-001s. The only common area between the two is the power and the emergency power-off switch.

Supports 256 devices

Each storage control in the 9390-002 supports up to 128 devices. Together they provide support for up to 256 devices in total.

Field upgrade from 9390-001 to 9390-002

A 9390-001 storage control connected to a single RAMAC 3 Storage Frame that supports 128 device addresses can be upgraded to a 9390-002 storage control that provides support for two RAMAC 3 Storage Frames and 256 device addresses.

Cache

The cache sizes are the same as those listed under “Cache” on page 76. The first storage control in the 9390-002 is configured like a 9390-001 with the same feature codes. You choose from the same size options in the second storage control but a different set of feature codes. (See “9390 Cache and NVS Feature Codes” on page 122 for a full list of 9390 feature codes.)

Nonvolatile storage

NVS is configured in a similar way to cache. The minimum is 2 x 32 MB and the maximum 2 x 128 MB. Any combination from the list under “Nonvolatile storage” on page 77 is valid. Each storage control in the 9390-002 has a different set of NVS feature codes.

Channels

Channels are configured as two sets of channels selected from the list on “Channels” on page 77, with each storage control having its own set of channels with unique feature codes.

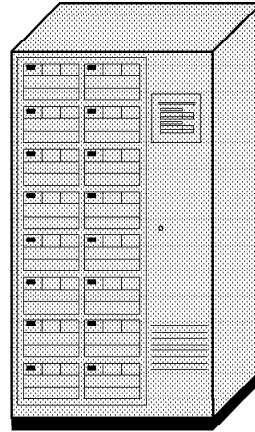
Common power supplies

The common power supplies power both storage controls. Foil “9390-001 Power Boundaries” on page 116 shows a schematic of the power connections and the power boundaries. There are two power cords for the 9390-002, each connecting to both storage controls of the 9390-002.

RAMAC 3 Storage Frame



- **New frame - 9391-A30**
 - Supports 128 logical volumes
 - Faster DDC interface
 - HSDA cards
- **9392-B33 drawer only**
 - No drawer intermix
- **Minimum 2 drawers**
 - 45.4 GB
- **Maximum 16 drawers**
 - 363.2 GB
- **Attaches to 9390 or 3990-6**



© IBM Corporation 1996

RAMAC 3 Storage Frame

The RAMAC 3 Storage Frame contains the RAMAC 3 drawers and connects to the 9390 and 3990-6 storage controls through the DDC interface. The RAMAC 3 Storage Frame also contains the HSDA cards that connect to the PDAD cards in each RAMAC 3 drawer.

New frame-9391-A30

The frame is externally the same as the existing 9391-A10 frame used by RAMAC and RAMAC 2 but has new drawer interfaces, HSDA cards, and internal cabling. The HSDAs provide support for 128 logical volume addressing. A 9391-A10 cannot be upgraded to a 9391-A30.

9392-B33 drawer only

The RAMAC 3 frame supports RAMAC 3 drawers only. There is no intermix with other RAMAC or RAMAC 2 drawers.

Minimum 2 drawers

The minimum configuration is two drawers supporting 16 logical 3390-3 volumes, for a capacity of 45.4 GB.

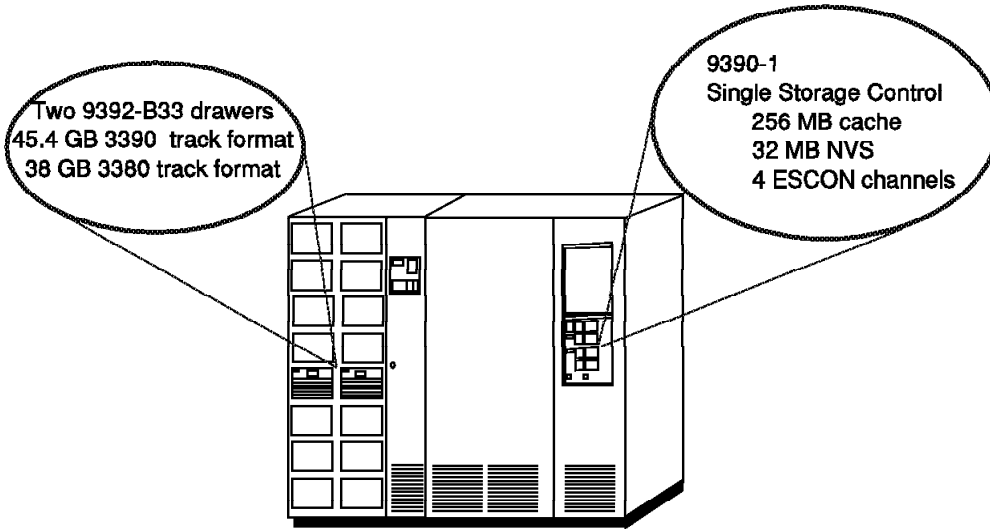
Maximum 16 drawers

The maximum configuration is 16 drawers, each emulating 8 logical volumes. This is a full 128 volume frame with a total capacity of 363.2 GB.

Attaches to 9390 or 3990-6

The RAMAC 3 Storage Frame attaches to either the 9390-001, 9390-002, or 3990-6, protecting your investment in 3990-6 technology. A 3990-6 and 9390-001 can attach to one RAMAC 3 frame. The 9390-002 attaches to two RAMAC 3 frames—one to each of the two storage controls.

Minimum Configuration: 45.4 GB



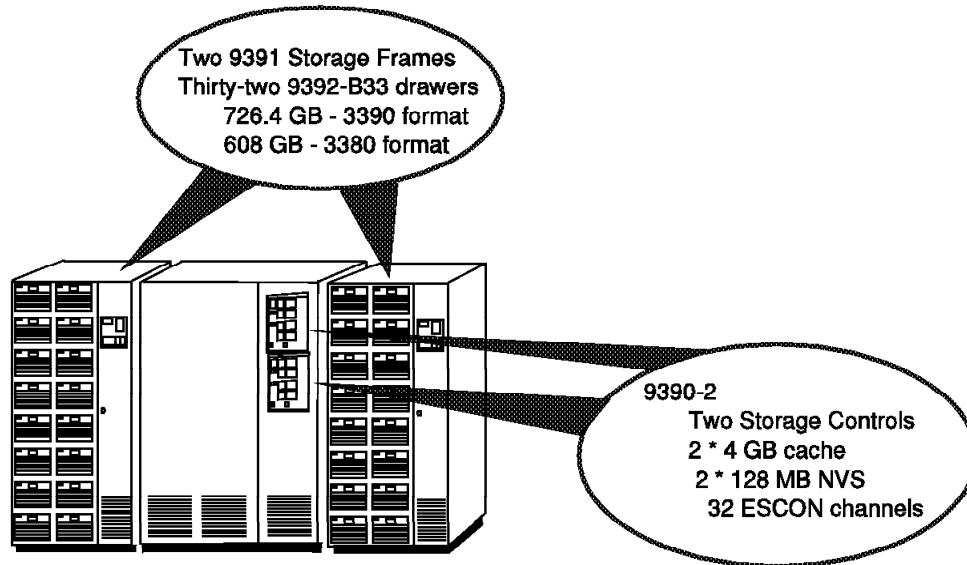
© IBM Corporation 1996

Minimum Configuration: 45.4 GB

This foil shows a minimum configuration of the RAMAC 3 Array Storage subsystem—using one 9390-001 and one RAMAC 3 Storage Frame. Note that even with the minimum number of two drawers, you still have a cache size of 256 MB and NVS size of 32 MB.

If the sixteen volumes in the minimum configuration are formatted in 3380 track format, the total capacity is reduced to 38 GB. In 3380 track format, each 3380 track is only 47476 bytes compared with 56664 bytes in 3390 format. The number of cylinders is the same (3339) in both formats.

Maximum Configuration: 726.4 GB



© IBM Corporation 1996

Maximum Configuration: 726.4 GB

This foil shows a maximum configuration of the RAMAC 3 Array Storage subsystem—using two RAMAC 3 Storage Frames and a 9390-002 Storage Control. In this configuration you would have 256 volumes in either 3380 or 3390 mode or any combination of 3380 and 3390. If all volumes are in 3390-3 format, the total capacity is 726.4 GB. If all volumes are in 3380 track format, the total capacity is 608 GB.

RAMAC 3 Addressing



- 9390-1 and 3990-6 can be set up to support:
 - 64 addresses
 - ▶ *One half-full frame of RAMAC 3*
 - 128 addresses
 - ▶ *One full frame of RAMAC 3*



© IBM Corporation 1996

RAMAC 3 Addressing

RAMAC addressing changes with RAMAC 3 because we now have up to 128 addresses in each subsystem.

9390 and 3990-6 can be set up to support:

Both the 9390 and 3990-6 support 128 addresses when attaching RAMAC 3 Storage Frames. However, if only 64 or fewer volumes are needed, both storage controls can be set up to address just 64—this saves having to define 128 UCBs when you plan to use only up to 64. (The ability to define fewer addresses than the storage control upper limit is similar to existing addressing options on a 3990-6, where you can define just 32 addresses on a 3990-6 that can support a maximum of 64).

If you limit the addressing to 64 volumes, you can install a RAMAC 3 Storage Frame with only a maximum of 8 drawers. If you define the full 128 addresses, you can install all 16 drawers.

RAMAC 3 Addressing: 64 Devices

IBM
San Jose
© IBM Corporation 1996

RAMAC 3 Storage Frame - half-full frame of 8 drawers

String 0

1																	2
3																	4
5																	6
7																	8
9	00	01	10	11	20	21	30	31	02	03	12	13	22	23	32	33	10
11	04	05	14	15	24	25	34	35	06	07	16	17	26	27	36	37	12
13	08	09	18	19	28	29	38	39	0A	0B	1A	1B	2A	2B	3A	3B	14
15	0C	0D	1C	1D	2C	2D	3C	3D	0E	0F	1E	1F	2E	2F	3E	3F	16

9390 or 3990-6 with 64 device addressing

AA BB CC DD EE FF GG HH ← Device address (logical)



© IBM Corporation 1996

RAMAC 3 Addressing: 64 Devices

This foil shows the addresses for a RAMAC 3 Storage Frame attached as string 0 with a maximum of eight drawers and the addressing limit set to 64. Note that the drawers are installed in the normal RAMAC sequence starting with drawer 9 and going down to drawer 16.

If you are installing this configuration and set the address limit to 64, only 64 devices have to be defined to the operating system and the input/output configuration program (IOCP).

With eight drawers, all physical device addresses are contiguous from x'00' to x'3F'. Note that the addresses themselves are not contiguous within a drawer, so, for example, drawer 9 contains addresses 00, 01, 10, 11, 20, 21, 30, and 31. Therefore if you do not have all eight drawers installed, you will see gaps in your addressing range.

The addressing scheme is developed from the current RAMAC and RAMAC 2 schemes. RAMAC 2 device addresses always occupy the same drawer location in RAMAC 3 configurations—the additional RAMAC 3 addresses are “overlaid” on the RAMAC 2 addresses.

Check with your IBM representative to discuss additional installation considerations before implementing the half-frame RAMAC 3 configuration.

RAMAC 3 Addressing: 128 Devices

IBM
CORPORATION
SAN JOSE
CALIFORNIA
95120

RAMAC 3 Storage Frame - full frame of 16 drawers

String 0

1	4E 4F 5E 5F 6E 6F 7E 7F	4C 4D 5C 5D 6C 6D 7C 7D	2
3	4A 4B 5A 5B 6A 6B 7A 7B	48 49 58 59 68 69 78 79	4
5	46 47 56 57 66 67 76 77	44 45 54 55 64 65 74 75	6
7	42 43 52 53 62 63 72 73	40 41 50 51 60 61 70 71	8
9	00 01 10 11 20 21 30 31	02 03 12 13 22 23 32 33	10
11	04 05 14 15 24 25 34 35	06 07 16 17 26 27 36 37	12
13	08 09 18 19 28 29 38 39	0A 0B 1A 1B 2A 2B 3A 3B	14
15	0C 0D 1C 1D 2C 2D 3C 3D	0E 0F 1E 1F 2E 2F 3E 3F	16

3990-6 or 9390 with 128 device addressing

AA BB CC DD EE FF GG HH ← Device address (logical)



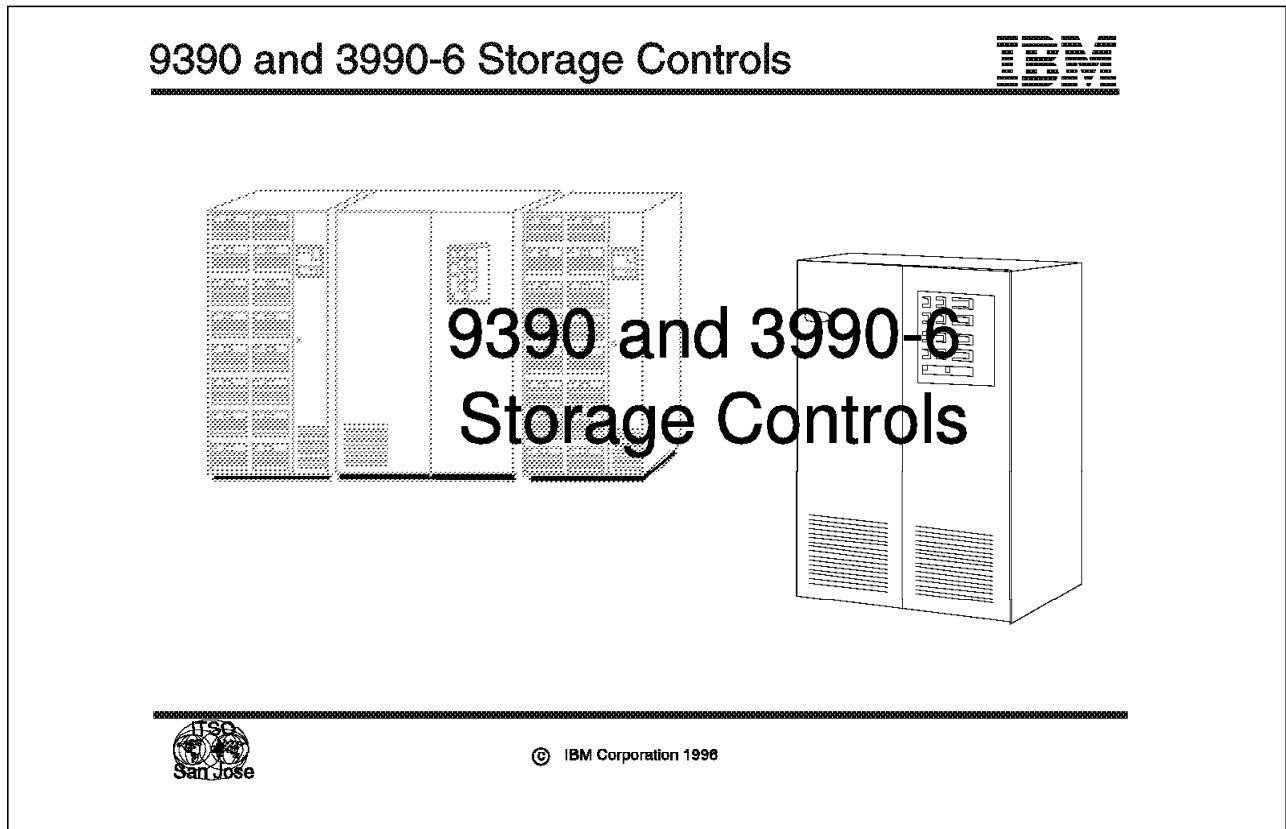
© IBM Corporation 1996

RAMAC 3 Addressing: 128 Devices

This foil shows the addressing scheme of a full 128-device frame attached to a 9390 or 3990-6 as string 0. You must define all 128 devices to your host system even if you have fewer than 128 devices installed.

The addressing is logically split into two parts, the lower eight drawers and the upper eight drawers. The lower eight drawers follow the same addressing convention as the 64 device addressing on foil "RAMAC 3 Addressing: 64 Devices" on page 85. The upper eight drawers follow a similar pattern starting at x'40' in drawer 8. This addressing scheme enables a concurrent transition from a half-frame of 64 addresses to a full-frame of 128 addresses.

Chapter 6. 9390 and 3990-6 Storage Controls



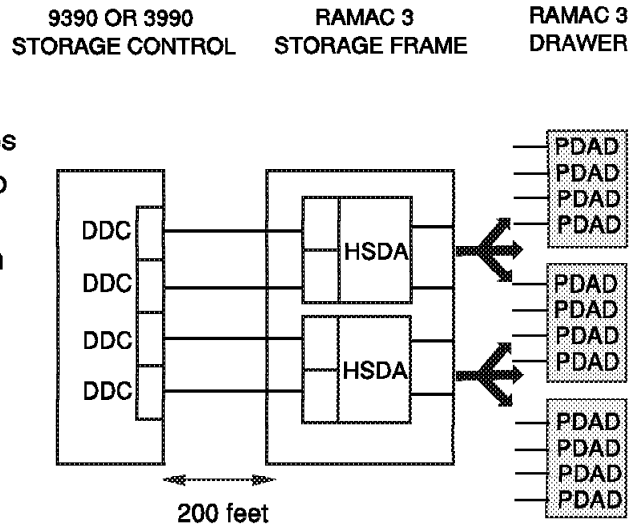
9390 and 3990-6 Storage Controls

In this chapter we describe enhancements to the storage control function provided by the 9390 and 3990-6 when attached to the RAMAC 3 Storage Frame.

High-Speed Device Adapter



- HSDA cards
 - Second-generation device adapter
 - Support for 128 devices
 - Data transfer rate of up to 11.9 MB/s
 - DDC-PDAD translation
 - Multipath "fan-out"
 - CE panel support
 - LIC load and activate
 - VPD
- DDC cable
 - New DDC cable for 9391-A30
 - Maximum length 200 ft



© IBM Corporation 1996

High-Speed Device Adapter

A device adapter (DA) is the interface between the storage control and the DASD (RAMAC or 3390). There are four DAs in a RAMAC rack or frame, arranged in pairs on two DA cards. The storage control is connected to the DAs by four DDC cables. These cables allow the storage control to be located up to 200 ft from the RAMAC rack or frame. Each DA provides one of the four paths to every drawer in a RAMAC rack or frame.

HSDA cards

RAMAC 3 introduces a second generation of device adapter, the high-speed device adapter (HSDA). The new adapter has a simplified design that supports data rates of up to 11.9 MB/s. In current adapters, the microprocessors in the adapter are involved in the setup and transfer of all data bytes. With the new simplified design, a microprocessor sets up the transfer and then passes control to a specialized hardware component that directly controls data flow through the adapter at high speed. The effective RAMAC 3 Array data transfer rate of the lower device interface is tripled compared with RAMAC or RAMAC 2.

There are two logic cards, each with two HSDA paths. The HSDA is responsible for data transfer, ECC, converting the DDC commands to PDAD commands, and directing data to each drawer. The HSDA provides the multiplexing, multipathing, and connection redundancy function from the 4 3990 or 9390 DDC channels to the 16 independent drawers in the RAMAC 3 drawer. The HSDA also provides protocol conversion from the DDC control unit protocol to the

PDAD drawer protocol and a level of drawer error correction also implemented on RAMAC and RAMAC 2 device adapters.

The HSDAs are configured through VPD settings and execute their own microcode. The HSDAs interface to the customer engineer (CE) panel and act as the “gateway” to the storage subsystem for all CE activity and diagnostics.

There are 128 logical devices supported in either 3390-3 or 3380 track format through the new DDC interface.

DDC cable

Four new DDC cables are required to connect the RAMAC Storage Frame to the 9390 or 3990-6 storage control. The new cables are needed to support the higher data rate and are of a higher specification than the standard DDC cables. The maximum length of the cable is 200 ft. Current DDC cables cannot be used with RAMAC 3.

Effects of High-Speed Device Adapter



- Faster destage from subsystem cache to drawer cache
- Faster stage from drawer cache to subsystem cache
- Faster PPRC initial volume synchronization
- Faster XRC initial volume synchronization
- Faster concurrent copy session management
- New and faster storage frame maintenance through FPGA technology
- No changes to storage control hardware

ENABLES COST EFFECTIVE, COMPACT, HIGH CAPACITY
RAMAC 3 ARRAY SUBSYSTEMS BY NEUTRALIZING THE IMPACT OF
HIGHER THROUGHPUT ON DRAWER PERFORMANCE AND
CONTROL UNIT EXTENDED FUNCTIONS



© IBM Corporation 1996

Effects of High-Speed Device Adapter

The faster HSDA improves all data transfers between the storage control (3990-6 and 9390) and the RAMAC 3 drawer. The RAMAC 3 drawer emulates eight logical volumes and might have to satisfy the higher I/O rates generated to this larger data pool. Therefore, the influence of the faster HSDA card may not surface in improved service times, but rather in preserving performance levels while sustaining higher throughput. Preserving or maintaining current levels of performance (say, those experienced with RAMAC 2 before consolidation and migration to RAMAC 3) at higher I/O rates might result in "go faster" changes to the 3990-6, 9390, RAMAC drawer, and Ultrastar 2XP disk drives being masked at the response time measurement level—the influence these changes have on I/O processing that improves the effectiveness and efficiency of RAMAC 3 Array Storage configurations merit attention and are mentioned here.

Faster destage from subsystem cache to drawer cache

Destaging operations are essential in any storage subsystem that implements write caching to nonvolatile storage. Nonvolatile storage is invariably a limited resource that benefits greatly from friendly and effective destage management techniques that use the "surrounding" hardware and interfaces. Writing data from the 3990-6 or 9390 cache to the RAMAC 3 drawer cache uses the lower interface components that largely govern the speed at which the destage operations take place. The HSDA cards in the RAMAC Storage Frame are the lower interface "bridge" between the storage control and the RAMAC 3 drawer. The HSDA cards enable destage operations to transfer data at triple the speed of RAMAC and RAMAC 2 storage subsystems. This is a major improvement for

RAMAC 3 Array Storage subsystems and paves the way for compact, high-performance, high-capacity configurations that exploit the 9.22 GB Ultrastar 2XP disk drive

Faster stage from drawer cache to subsystem cache

The HSDA facilitates faster data transfer from subsystem cache to drawer cache and from drawer cache to subsystem cache. The cache bandwidth of the 3990-6 and 9390 and their storage path processor capabilities support the higher transfer speeds. The RAMAC 3 drawer also supports the HSDA higher data transfer capability through the faster PDAD interface. All of this results in faster data transfers for both staging and destaging operations.

Faster PPRC initial volume synchronization

In a PPRC environment, primary volume contents are first copied to the secondary volume, and subsequent updates are written to both volumes. The primary and secondary volumes are considered a PPRC pair when their contents are identical. The subsequent record updates to the primary volume and the corresponding shadowed record update to the secondary volume do not use the lower interface components—instead, data is written to cache and NVS in both primary and secondary volume storage controls. However, the initial volume copy, which entails reading data from the primary and writing it to the secondary volume, uses the lower interface components. Therefore the faster HSDA speeds up the PPRC volume synchronization process by supporting data transfer operations from the RAMAC 3 drawer to the storage control of up to 11.9 MB/s. Of course, the time taken to synchronize PPRC primary and secondary volumes depends on other storage control and RAMAC 3 drawer workloads.

Faster XRC initial volume synchronization

XRC requires the primary and secondary volume contents to be identical when they are included in an XRC session. Similar to PPRC, primary volume contents are first copied to the secondary volume, and subsequent updates are written to both volumes. The initial process of synchronizing the two volumes requires data transfer from the physical primary device to the physical secondary device. When the primary and secondary devices are RAMAC 3 drawers, the faster HSDA enables transfers of up to 11.9 MB/s.

Faster concurrent copy session management

The 3990-6 and 9390 concurrent copy function enables customers to take point-in-time copies of data by using a cache sidefile that is loaded with the time-zero version of the data. Time-zero refers to the state of the data when the concurrent copy session is started and before it gets updated. Before a record is updated, it is copied to the cache sidefile, thus creating a “before update” version or time-zero version of the record. After the cache sidefile data has been prepared, the record can be updated on the physical device—the physical device now contains the updated version while the cache sidefile contains the time-zero version of the data. Concurrent copy relies on effective and efficient sidefile management. The time taken to “load” the cache sidefile is improved by the faster HSDA used in RAMAC 3 drawer configurations because the staging operation can take place at 11.9 MB/s.

New and faster storage frame maintenance through FPGA technology

The faster HSDA cards are located in the RAMAC Storage Frame. They connect to the DDC interface in the control unit and the PDAD interface in the RAMAC 3 drawers. The HSDA uses a new technology called *field programmable gate arrays* (FPGAs) which enable hardware logic to be upgraded concurrently in the field. This is a more effective and efficient way of making functional enhancements, corrections, or changes to storage subsystem components than programming microcode workarounds. The use of FPGA technology offers a new and faster concurrent maintenance platform for the RAMAC 3 Storage Frame and addresses the continuing demand for planned outage avoidance techniques that satisfy continuous availability requirements.

No Changes to Storage Control hardware

The 3990-6 and 9390 storage controls do not have to be changed to take advantage of the new HSDA. The 3990-6 and 9390 hardware components already support the higher data transfer speeds now attainable through the faster HSDA for RAMAC 3 Array Storage configurations. (Some early models of the 3990-6 require their internal hard drive to be upgraded).

Sequential Detect



- Sequential detect is available on both the 3990 Model 6 and 9390
- Sequential detect is applicable to:
 - Read data
 - Data that is cachable
- Sequential detect is invoked if more than three cylinders are read sequentially
- Records are left in cache for normal LRU processing to allow reuse
- VPD option to turn off: default is ON



© IBM Corporation 1996

Sequential Detect

Sequential detect is a new 3990-6 and 9390 cache algorithm that complements the software sequential staging requests issued by the various host access methods.

Sequential detect is available on both the 3990-6 and 9390

Sequential detect is a function of the 9390 and 3990-6 that recognizes sequential access patterns even when the host software does not set the sequential attribute in the channel program. When the 9390 or 3990-6 determines the sequential nature of the workload, it initiates sequential prestaging of data from the RAMAC 3 drawer. The sequential prestaging enables multiple tracks to be staged to drawer and subsystem cache in anticipation of use by the application.

Access methods where the access pattern can be both random and sequential do not set the sequential attribute in the channel program, and the data is handled by the storage control using its nonsequential caching algorithms. The new algorithm attempts to detect a sequential reference pattern, and, if the application continues to access sequentially, the storage control invokes sequential staging. The storage control can stage up to 15 tracks at one time, substantially improving the effective throughput of the application.

Sequential detect is applicable to:

The new algorithm applies to read data that is cachable. Data that uses bypass cache or ICL indicators can take advantage of sequential detect when the bypass cache and ICL intent is ignored through changing default VPD settings. Applications using VSAM key-sequenced data sets (KSDSs) can take advantage of this change as VSAM does not set the sequential attribute when processing data sequentially. Non-IBM software that does not set the sequential intent indicator but accesses data sequentially will also benefit from sequential detect.

Sequential detect is invoked if more than three cylinders are read sequentially

The current sequential detect algorithm is not invoked unless the application reads more than three cylinders of data from the same data set sequentially. The algorithm does this to ensure that the application can take advantage of the cylinder staging that will be invoked when the sequential staging starts.

Records are left in cache for normal LRU processing to allow reuse

The main difference between access-method-initiated sequential staging and the sequential detect algorithm is the status of the staged tracks after they have been read by the application. The access-method-initiated sequential algorithm makes the old tracks except the last *least recently used*. This ensures that the cache segments the tracks occupy will be reused as soon as the storage control needs more cache space. The intention is to stop sequential activity from flooding the cache with sequential data to the detriment of other, often more critical, online data.

The sequential detect algorithm takes a different view: If the data was really processed only sequentially, the access method would have used the sequential attribute in the channel program. The assumption is that the data may be read sequentially and randomly—the data is left for the LRU algorithm to handle.

Sequential detect also complements the faster HSDA and drawer processing that facilitates up to three times the data rate when compared with RAMAC, and RAMAC 2 Array DASD.

VPD option to turn off: default is ON

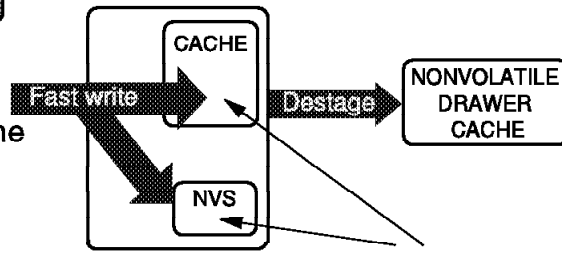
As this option brings benefits to all types of workloads, the decision has been made to make ON the default for sequential detect. Sequential detect can be set to OFF at the service panel.

Enhanced NVS Management for RAMAC 3



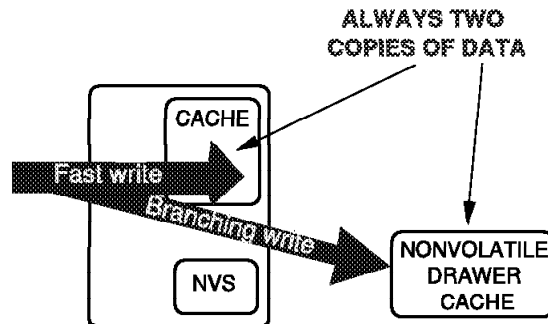
Current 3990-6 NVS management

- Data written to cache and NVS
- DE issued when data in cache and NVS
- Asynchronous destage to drawer cache



Enhanced NVS management

- Data written to subsystem cache and drawer cache
- DE issued when data written to subsystem cache and drawer cache



© IBM Corporation 1996

Enhanced NVS Management

The 9390 and 3990-6 implement a caching algorithm change that improves sequential write performance.

Current 3990-6 NVS management

The traditional NVS management algorithm uses the 3990 NVS to hold a second copy of all modified data. As cache is volatile and its copy of data is lost if you have a power failure, the NVS copy, protected by a battery, is used to write to disk once power is restored. The I/O is considered complete when a copy of the data is written to both cache and NVS and channel end (CE) and device end (DE) are issued to the host system. When space is needed in cache or NVS, the data is destaged to the physical disk—the writing to disk thus takes place asynchronously with the posting of I/O completion status.

The traditional algorithm works well for most types of workloads. However there are occasions where a heavy sequential workload mixed with an online workload can give less than optimum performance. The sequential workload rapidly fills NVS and causes the storage control into a forced destage operation to free up space in NVS. This can have unpredictable results on your online response time. Optional changes announced in March 1996 are designed to help alleviate this effect, particularly for DB2 workloads, by changing the freespace thresholds in NVS.

Enhanced NVS management

This new NVS management change takes advantage of the higher-speed interface between the storage control and the RAMAC 3 Storage Frame. It uses the nonvolatile storage in every RAMAC 3 drawer rather than the storage control to hold sequential modified data. For write operations, the first copy of the data is still written to subsystem cache, but the second copy is now written to drawer cache instead of subsystem NVS. The I/O is signaled complete (DE issued to the host system) when the second copy of data is in drawer cache—drawer cache is nonvolatile because its contents are protected by the drawer battery.

The advantages of this technique are to remove large blocks of sequential data from NVS and allow NVS to be used primarily for online DASD fast write (DFW) and record cache hits. Mixing a batch workload with an online workload also becomes more practical. (In the past, with small amounts of storage behind a control unit, it was possible to separate online and batch onto different controllers. Now with up to 363.2 GB behind a single storage control, this separation becomes less practical, and the need to be able to handle both online and batch on the same storage control is more important.)

The new algorithm selects data to send straight to the drawer cache according to the expected transfer size (taken from information in the channel program). If the number of records, or the transfer size is large enough, the algorithm directs output to the drawer cache and at the same time keeps a copy in the cache. The algorithm can also be triggered if NVS is filling up. As the high threshold is reached, all sequential writes are directed to the drawer cache.

The effect of this new algorithm is to reduce the probability of the storage control NVS becoming full and thus avoid the unpredictable impact that NVS-full has on performance. The new NVS management algorithm retains the highest levels of data availability by always ensuring that there are two copies of data at all times.

Enhanced NVS Management...



- Takes advantage of faster lower interface
- Exploited by Sequential Detect
- Uses nonvolatile drawer cache for second copy
- Designed for long chain or large block size sequential data streams
- Reduces NVS-full probability
- Avoids asynchronous destage to drawer
- Improved performance for mixed batch and online
- Used for other writes if NVS becomes full

**PROVIDES 64 MB OF NONVOLATILE STORAGE FOR
QUALIFYING I/O TO RAMAC 3 LOGICAL VOLUMES**



© IBM Corporation 1996

Enhanced NVS Management...

Takes advantage of faster lower interface

The new NVS management algorithm takes advantage of the higher-speed lower interface in the RAMAC 3 Array Storage or a 3990-6 with RAMAC 3. The lower interface speed of the RAMAC 3 Array Storage has tripled to 11.9 MB/s, enabling fast data transfer between the storage control and the RAMAC 3 drawer in both directions. The increase in lower interface speed makes it possible to "bypass" subsystem NVS and write straight to the drawer cache.

Exploited by sequential detect

The new NVS management algorithm exploits the faster lower interface capability of RAMAC 3 Array Storage by writing updated data straight to nonvolatile drawer cache without writing a copy to NVS. The bypassing of NVS is of particular significance for sequential workloads where a large amount of update activity could result in filling NVS and invoking threshold management tasks that could impact application performance. The 9390 and 3990-6 storage controls enable the new NVS management algorithm when the host software signals sequential intent through the define extent (DX) channel command. The scope of the new NVS management technique is broadened substantially when the storage control determines that an application is using sequential access even though that is not indicated in the DX command. Sequential detect capability in the 9390 and 3990-6 ensures that the lower-interface enhancements combine with the new NVS management algorithms to improve throughput and response times.

Uses nonvolatile drawer cache for second copy

Drawer cache is also nonvolatile and provides a dedicated NVS for exclusive use by the eight logical volumes. The data written to drawer cache becomes subject to the RAMAC 3 drawer's advanced destage management techniques, which take account of drawer cache occupancy, disk drive utilization, and RAID-5 parity processing.

Designed for long chain or large block size sequential data streams

The new algorithm is activated when a large write transfer size is detected or a large number of chained records is detected. When a write transfer of this nature is detected, data is transferred to cache and RAMAC 3 drawer cache in a branching write fashion similar to the operation used when data is transferred to both cache and channel in traditional 3990 configurations.

Reduces NVS-full probability

Because large sequential write operations now bypass the storage control NVS, more space is available for online data, and there is less probability that NVS will become full because of a burst of write activity. The sequential data is now stored in "local" drawer cache where drawer destage management ensures effective resource utilization while controlling RAID-5 requirements.

Avoids asynchronous destage to drawer

The current storage control algorithms maintain two copies of the data—one copy in cache and the other in NVS. The new NVS management algorithm also maintains two copies of the data—one in cache and the other in nonvolatile drawer cache. With the current implementation, an asynchronous destage of the data from subsystem cache to drawer cache is performed as data ages and has to be secured on the "device"—in this case, the RAMAC 3 drawer. The new NVS management algorithm writes qualifying data directly to the RAMAC 3 drawer thus avoiding the asynchronous destage at a later time.

Improved performance for mixed batch and online

With larger configurations you are less able to split the batch workload from the online workload and less able to manage the interaction between the two. This NVS management change directs large sequential writes away from the online I/O.

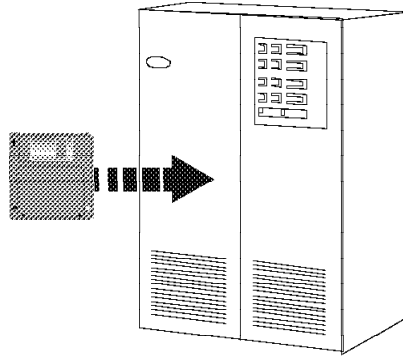
Used for other writes if NVS becomes full

The algorithm cuts in at various levels to direct other writes to drawer cache as NVS becomes full.

RAMAC 3 LIC Download and Activation



- 3990-6 LIC upgrade is disruptive
 - Provides support for 128 addresses
 - Optimized for faster device adapter
- 9390-001 and 9390-002
 - Full RAMAC 3 support shipped with box
- 9390-002
 - Each storage control can be updated independently



© IBM Corporation 1996

RAMAC 3 LIC Download and Activation

The 3990 has increased its functional capability many times over many years. The functional additions and enhancements are predominantly delivered through LIC. Thus the loading and activation of LIC must always be considered when DASD subsystems graduate to the latest levels of storage control functionality.

3990-6 LIC upgrade is disruptive

The 3990-6 microcode upgrade to support RAMAC 3 is disruptive. However, as RAMAC 3 cannot be attached with any other DASD, this should not be a problem. Data must be migrated from an existing 3990 configuration before RAMAC 3 is attached. The new LIC level provides support for 128 addresses.

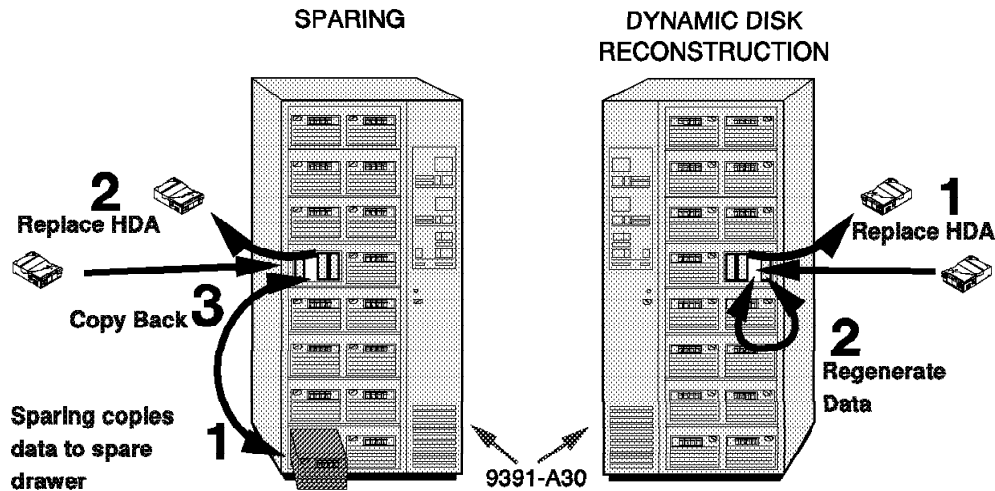
9390-001 and 9390-002

A 9390 will be delivered with the RAMAC 3 LIC already installed as RAMAC 3 is the only device supported by the 9390.

9390-2

Each storage control in the 9390-2 is independent. Should LIC upgrades be required in the future, each storage control can be updated when convenient for the customer.

RAMAC 3 Drawer Maintenance



Concurrent Access to Data during Drawer Maintenance

SELECT THE MAINTENANCE OPTION THAT SUITS YOUR INSTALLATION BEST



© IBM Corporation 1996

RAMAC 3 Drawer Maintenance

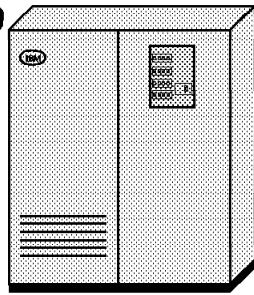
A key requirement of storage subsystems today is continuous operations. To address this requirement the RAMAC Array DASD subsystem introduced the concept of concurrent drawer maintenance with the sparing function. RAMAC 2 Array DASD enhanced the concurrent drawer maintenance philosophy further and broadened the maintenance options by introducing dynamic disk reconstruction. RAMAC 3 Array Storage and the 3990-6 attaching to the RAMAC 3 drawer retain the strong focus on concurrent drawer maintenance by supporting both Sparing and dynamic disk reconstruction.

With a larger amount of data now stored on the disk drives in the RAMAC 3 drawer, it is imperative to offer as flexible and broad a concurrent maintenance package as possible to avoid planned outages.

RAMAC 3 Sparing



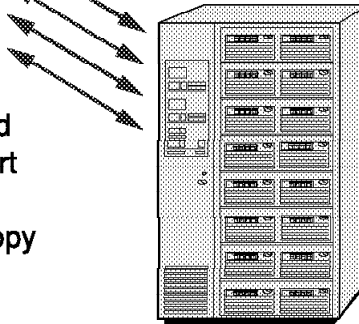
3990-6 or 9390



- Choice of two sparing options
 - Dynamic
 - Manual
- Choice of three post repair options
 - Immediate copyback
 - Delayed copyback
 - Floating spare

- Enhanced performance
- One or two spare drawers
- Application performance prioritized
- Comprehensive monitoring support
 - SIM reporting
- Data redundancy restored after copy
- Enables concurrent repair for all drawer components

9391-A30



© IBM Corporation 1996

RAMAC 3 Sparing

There are two sparing options—Dynamic Sparing and manual sparing.

Dynamic Sparing has the advantage of initiating automatically on fault detection, thus providing maximum data loss protection by copying the data to the spare drawer and restoring RAID-5 redundancy as soon as possible. Manual sparing offers customers the flexibility of invoking sparing for specific drawers, and provides the choice of not reserving a spare drawer, but instead, of quiescing a drawer when appropriate and initiating the sparing process with the help of a service representative.

Sparing with a RAMAC 3 configuration is similar to the process we are familiar with on RAMAC and RAMAC 2 subsystems. The default for RAMAC, RAMAC 2, and RAMAC 3 is to automatically start the sparing process, assuming that a spare has been defined to the subsystem. Dynamic disk reconstruction is used if a spare drawer is not assigned and the component malfunction is a drive failure.

The expected time for a sparing operation will be the same as for RAMAC 2 even though the RAMAC 3 drawer must now protect eight logical volumes as opposed to four and restore full RAID-5 capability to the target drawer. Sparing using RAMAC 3 drawers takes advantage of the faster device interfaces to the storage control and enhances the staging and destaging process during drawer copy.

During the sparing process, application performance is given priority by interleaving the copying with "favored" host access—the host is allowed to

interrupt the copying process so that application performance is not affected too much. SIMs are generated and sent to the host during the Dynamic Sparing process so that operations staff are aware of ongoing status. Up to two drawers can be defined as spare drawers, but only one can be the target of the sparing operation at any time. We recommend one spare drawer per RAMAC 3 Storage Frame.

Sparing is the only drawer maintenance function that protects against all drawer failures—dynamic disk reconstruction should be used only for disk drive module (DDM) failures.

When the failed drawer has been repaired, the data can be copied back immediately, copied back after a specified delay, or remain on the copied-to drawer. These three options are called *immediate copyback*, *delayed copyback*, and *floating spare environment*, respectively.

RAMAC 3 Drawer Maintenance - Summary

FLEXIBILITY OF TWO CONCURRENT REPAIR OFFERINGS

- **Dynamic Sparing preferred**
 - Provides concurrent repair for all drawer components
 - Restores RAID-5 capability to target drawer
 - Initiates automatically on detection of first component outage
 - Customer plans for time to copy data to spare drawer
- **Dynamic disk reconstruction**
 - Avoids spare drawer configurations
 - Avoids time to copy data before repair
 - Enables concurrent repair for disk drive only

RAMAC 3 ALLOWS CUSTOMER TO CHOOSE AT INSTALLATION TIME



© IBM Corporation 1996

RAMAC 3 Drawer Maintenance - Summary

Not every storage subsystem function or solution is globally adopted by all customers. The RAMAC 3 Array Storage (and RAMAC 2 Array DASD) recognizes that device maintenance is an element of storage subsystem management where repair alternatives differ among customers according to data availability and continuous operations requirements. Thus RAMAC 3 Array Storage offers a choice of drawer maintenance procedures.

Dynamic Sparing preferred

Dynamic sparing copies all of the data from the suspect drawer to a spare drawer. When all of the data is copied to the spare drawer, it becomes the active drawer by servicing all I/O to the data and thus "redirecting" I/O away from the suspect drawer. Thus any repair action can take place on the suspect drawer because it is now an inactive drawer. Copying all of the data to another drawer therefore enables concurrent repair for any drawer component with no impact on performance.

The spare drawer that the data is copied to is a fully functional enhanced RAID-5 RAMAC device that assumes active drawer status. Therefore when the data is copied to the spare drawer, it is immediately and automatically protected from data loss in the event of drawer component malfunction.

Dynamic Sparing is automatically invoked by the RAMAC 3 drawer and the storage control on detection of the first error that affects the level of redundancy

in the drawer. The customer does not have to be present or decide on which course of action to follow.

The RAMAC 3 drawer contains 22.7 GB of user capacity stored on eight logical volumes. Dynamic Sparing ensures that application access to the data receives priority during the copying, thus minimizing the impact on application performance. When the failing drawer has been repaired, the data can either be copied back immediately, copied back at a more convenient time, or retained on the new active drawer leaving the repaired drawer as the new spare. If the data is to be copied back to the repaired drawer, you therefore have the flexibility to plan for the most suitable timeframe.

Dynamic disk reconstruction

Dynamic disk reconstruction offers an alternative method of concurrent DDM maintenance. It applies to the DDM component of the drawer only and enables disks to be replaced concurrently with drawer I/O activity. While the DDM is being replaced, the enhanced RAID-5 function regenerates the data for the failed disk. After the DDM has been replaced, the data is reconstructed and written to the new disk.

Dynamic disk reconstruction does not offer the global “all-component” repair philosophy of Dynamic Sparing because the technique applies only to DDM components, but it does enable installations to consider a “no-spare” policy for DDM repair scenarios. Dynamic disk reconstruction also avoids the time taken to copy data to the spare drawer, but we must remember that Dynamic Sparing “paces” the copy to minimize the impact on application performance.

RAMAC 3 Array configurations offer a unique combination of drawer maintenance strategies. Dynamic disk reconstruction can be used for DDM repairs while spare drawers are assigned in the RAMAC 3 Storage Frame. The option of enabling dynamic disk reconstruction for DDM repairs while simultaneously enabling sparing, is selected at installation time—this option can be changed after installation to suit customer requirements.

Chapter 7. Installation Planning

RAMAC 3 Array Storage



Installation Planning and Data Migration



© IBM Corporation 1996

Successful implementation of RAMAC 3 Array Storage entails the execution of several planning tasks. These planning tasks include physical setup, applying software updates, loading and activating new levels of storage control LIC, and defining new device types to installation hardware and software systems. This chapter describes some of the major tasks to consider when planning for RAMAC 3 Array Storage implementation.

Installation and Migration



- Software dependencies
- IOCP definition
- Physical planning
- VPD settings
- ESCON connections
- 3990 Model 6 support for IBM RAMAC 3 Storage Frame
- Data migration



© IBM Corporation 1996

Installation and Migration

For the IBM RAMAC 3 Array Storage a system assurance product review should be conducted. The *Systems Assurance Process Review (SAPR) Guide*, SA95017-03 can be used to assist in the review. The guide is available through HONE SA ADVISOR. To view the guide on the SAPRDOC LEXVMIC1 tools disk, issue the following command:

```
TOOLS SENDTO LEXVMIC1 TOOLS SAPRDOC GET SA95017 LIST3820 *
```

Storage product specialists and customer engineers should review this document when ordering and planning the installation of the IBM RAMAC 3 Array Storage.

Several key aspects of installation planning are described on the foil.

Software dependencies

Check that the system control program (SCP) level supports the IBM RAMAC 3 Array Storage and install all necessary program temporary fixes (PTFs). The preventive service planning (PSP) records provide all of the information you need for installation planning. Call your software service center to get the PSP record entries you want.

See Chapter 10, "Software" on page 139 for details on software support for the RAMAC 3 Array Storage.

IOCP definitions

RAMAC 3 Array Storage supports up to 256 devices, which must be defined to the SCP and the channel subsystem. Similarly, the 3990-6, connected to the RAMAC 3 Storage Frame, supports up to 128 devices. The hardware configuration definition (HCD) and IOCP programs define devices to the channel subsystem. The devices, and the paths that they use, must be defined correctly to enable efficient application I/O processing. See “9390-001 IOCP Definition Example” on page 108 for a detailed description of sample IOCP definitions for the RAMAC Array Storage.

Physical planning

The *9391/9392 RAMAC Array Storage Reference Summary*, SX26-1682, gives detailed information about physical planning. Environmental data such as floor space and power requirements can be found in “Physical Specifications” on page 114.

VPD settings

RAMAC 3 Array Storage configurations allow dynamic disk reconstruction to take priority over Dynamic Sparing for DDM failures through a VPD setting. The VPD setting enables a new service panel option that performs this function.

ESCON connections

The maximum throughput capability of the IBM RAMAC 3 Array Storage is obtained when the 9390-001, 9390-002, or the 3990-6 connection uses 18 MB/s ESCON channels.

3990-6 support for IBM RAMAC 3 Array Storage

After the 3990-6 storage controls have been updated with the correct microcode level, they support the IBM RAMAC 3 Storage Frame.

Data migration

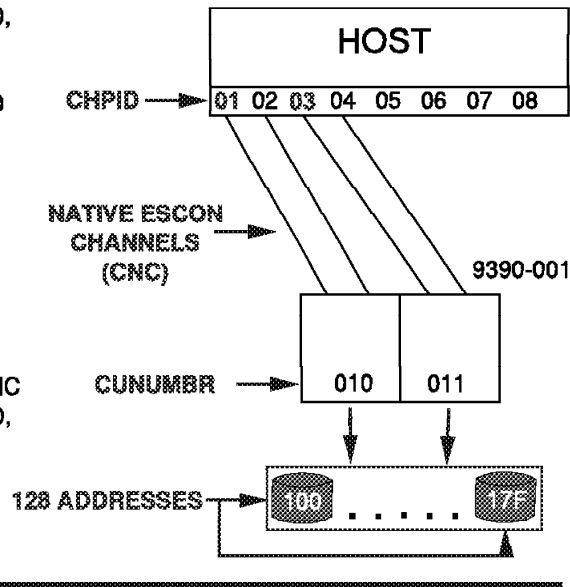
Plan data migration before installing the IBM RAMAC 3 Array Storage. Use peer-to-peer remote copy (PPRC), PPRC dynamic address switching (P/DAS), and extended remote copy (XRC) to copy the data from the old device to the IBM RAMAC 3 Array Storage. You must use DFSMS or DFDSS when you migrate data from different device types. The RAMAC Data Migrator can also be used to migrate to the RAMAC 3 Array Storage.

9390-001 IOCP Definition Example

```
CHPID PATH=((1),(2),(3),(4)),TYPE=CNC
CNTLUNIT CUNUMBR=010,UNIT=3990,
        UNITADD=((00,128)),
        PATH=(1,2)
CNTLUNIT CUNUMBR=011,UNIT=3990
        UNITADD=((00,128)),
        PATH=(3,4)
IODEVICE CUNUMBR=(010,011),
        UNIT=3390,UNITADD=00,
        ADDRESS=(100,128)
```

OR

```
CHPID PATH=((1),(2),(3),(4)),TYPE=CNC
CNTLUNIT CUNUMBR=010,UNIT=3990,
        UNITADD=((00,128)),
        PATH=(1,3,2,4)
IODEVICE CUNUMBR=(010),
        UNIT=3390,UNITADD=00,
        ADDRESS=(100,128)
```



© IBM Corporation 1996

9390-001 IOCP Definition Example

The foil shows the correct IOCP definitions for the RAMAC 3 Array Storage (9390-001). The only change is the 128 address capability. In the CNTLUNIT macro, set the UNITADD device count to 128, and in the IODEVICE macro, set the ADDRESS device count to 128.

In the diagram, CUNUMBR 010 is equivalent to 9390-001 cluster 0, and CUNUMBR 011 is equivalent to 9390-001 cluster 1. Cluster 0 is connected to processor channel paths 01 and 02, and cluster 1 is connected to processor channel paths 03 and 04. Notice that the number of CNTLUNIT macros used in the IOCP definition has an effect on the channel subsystem's path selection technique.

Two-CNTLUNIT macro definition

When two CNTLUNIT macros are used to define the paths to the 9390-001, the channel subsystem selects the first CHPID defined by the first PATH statement in the first CNTLUNIT macro—namely, CHPID 01 that attaches to CUNUMBR 010. Using this channel path, the channel subsystem attempts to acquire a path to the requested device. If the channel subsystem cannot obtain a path to the device it tries another path—but which one?

The channel subsystem's second choice is the first CHPID defined by the first PATH statement in the *second* CNTLUNIT macro—namely, CHPID 03 that attaches to CUNUMBR 011. This is the most efficient second choice because the path busy or unattainable status from the first attempt might have been due to both

storage paths in cluster 0 being busy or unavailable for some other reason. It therefore makes sense to attempt to establish a connection to the device by using a channel path connected to the other cluster—cluster 1.

The third choice path is the second CHPID defined by the first PATH statement in the first CNTLUNIT macro (CHPID 02 connected to CUNUMBR 010). The fourth choice of channel path is the second CHPID defined by the second PATH statement in the second CNTLUNIT macro (CHPID 04 connected to CUNUMBR 011).

The channel subsystem uses a rotational path selection mechanism based on the IOCP definition. The IOCP definition should be coded to enable the channel subsystem to “ping-pong” between storage control clusters on successive path selection attempts. Successive attempts to establish a connection to the device by using interfaces on the same cluster might delay I/O processing and data transfer.

Single-CNTLUNIT macro definition

When a single CNTLUNIT macro is used to define the paths to the 9390-001, the same logic applies. The foil shows the definition for a single CNTLUNIT macro definition for the same four-path configuration as before. Notice that the single PATH statement defines the sequence of CHPID selection as 1, 3, 2, 4. This has the same effect as the two-CNTLUNIT macro definition because, in the absence of the second CNTLUNIT macro, the channel subsystem cannot “ping-pong” between PATH statement CHPIDs. Instead, the channel subsystem attempts connection to the device by trying each of the CHPIDs *in the order in which they appear in the single PATH statement in the CNTLUNIT macro*.

In both cases, channel path connections to the 9390 and 3990-6 cluster interfaces must be well understood before coding the IOCP definitions.

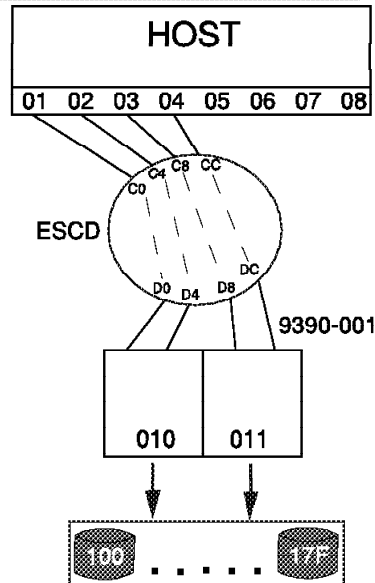
When only eight RAMAC 3 drawers are installed in the IBM RAMAC 3 Storage Frame, the number of devices can be set to 64 in the 9390 or 3990 VPD setting and in the IOCP definition statement.

9390-001/ESCD IOCP Definition Example



```
CHPID PATH=((1),(2),(3),(4)),TYPE=CNC,SWITCH=01
CNTLUNIT CUNUMBR=A00,UNIT=SWCH,
    UNITADD=((00,1)),
    PATH=(1,4)
LINK=(FE,FE)
IODEVICE CUNUMBR=A00,UNIT=SWCH,
    ADDRESS=A00
```

```
CNTLUNIT CUNUMBR=010,UNIT=3990,
    UNITADD=((00,128)),
    PATH=(1,2)
LINK=(D0,D4)
CNTLUNIT CUNUMBR=011,UNIT=3990
    UNITADD=((00,128)),
    PATH=(3,4)
LINK=(D8,DC)
IODEVICE CUNUMBR=(010,011),
    UNIT=3990,UNITADD=00,
    ADDRESS=(100,128)
```



© IBM Corporation 1996

9390-001/ESCD IOCP Definition Example

The foil shows the IOCP definitions for a 9390-001 connected to a host system through an ESCON Director (ESCD). We use the two-CNTLUNIT macro definition in our example. The channel path selection method described in the previous foil is used, and the only difference is the definition of the ESCD and how logical paths are defined through IOCP coding.

ESCON Director Definition

The ESCD must be defined to the channel subsystem by using the CHPID, CNTLUNIT, and IODEVICE macros as before and defining the ESCD as a switch through the SWCH value for the UNIT keyword on the CNTLUNIT and IODEVICE macro. An ESCD is identified by a switch number coded in the CHPID macro—the ESCD used in our example is coded SWITCH=01. The ESCD is accessed through CHPIDs 1 and 4 as specified by the PATH parameter of the SWCH control unit, and the channels used are native ESCON channels as specified by TYPE=CNC. The ESCD has a CUNUMBR of A00 and an ADDRESS of A00. The PATH parameter defines the CHPIDs that the host uses to communicate with the ESCD control ports—CHPIDS 01 and 04 both switch to ESCD port FE, enabling host software such as the ESCON Manager (ESCM) to monitor and alter ESCD port states.

9390/3990-6 Definition

The PATH parameter in the CNTLUNIT macro defines the CHPIDs used to access the devices through the appropriate cluster (CUNUMBR 010 and CUNUMBR 011). Below the PATH parameter is a LINK parameter that describes the ESCD port addresses used by the CHPIDs in the PATH statement. The PATH and LINK parameter values are positional—that is, CHPID 01 uses ESCD port address D0 and CHPID 02 uses ESCD port address D4 to access the devices behind the 9390 or 3990-6. Notice that the ESCD ports described by the LINK parameter are the “outbound” ports that connect the ESCD to the 9390 or 3990-6.

9390-002 IOCP Definition Example



CHPID PATH=((1),(2),(3),(4),(5),(6),(7),(8)),TYPE=CNC

CNTLUNIT CUNUMBR=010,UNIT=3990,
UNITADD=((00,128)),
PATH=(1,2)

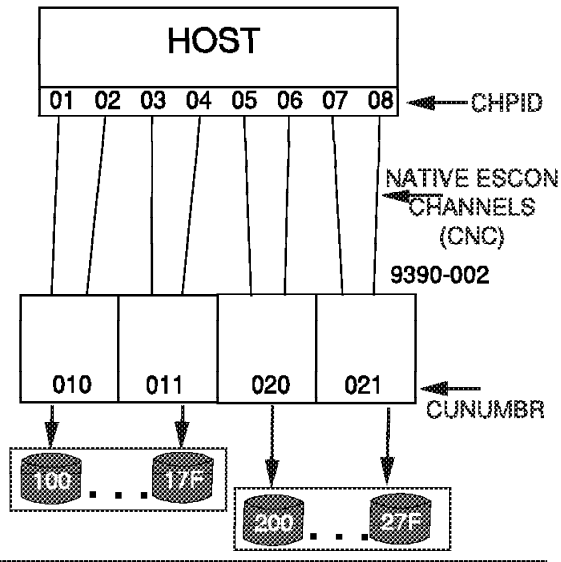
CNTLUNIT CUNUMBR=011,UNIT=3990
UNITADD=((00,128)),
PATH=(3,4)

CNTLUNIT CUNUMBR=020,UNIT=3990,
UNITADD=((00,128)),
PATH=(5,6)

CNTLUNIT CUNUMBR=021,UNIT=3990
UNITADD=((00,128)),
PATH=(7,8)

IODEVICE CUNUMBR=(010,011),
UNIT=3990,UNITADD=00,
ADDRESS=(100,128)

IODEVICE CUNUMBR=(020,021),
UNIT=3990,UNITADD=00,
ADDRESS=(200,128)



© IBM Corporation 1996

9390-002 IOCP Definition Example

The 9390-002 comprises two storage controls in a single frame, and the IOCP definitions can be considered similar to those required for two 9390-001 storage subsystems.

9390-002/ESCD IOCP Definition Example



CHPID PATH=((1),(2),(3),(4),(5),(6),(7),(8)),TYPE=CNC,SWITCH=01

CNTLUNIT CUNUMBR=010,UNIT=3990,
UNITADD=((00,128)),
PATH=(1,2)
LINK=(A0,A4)

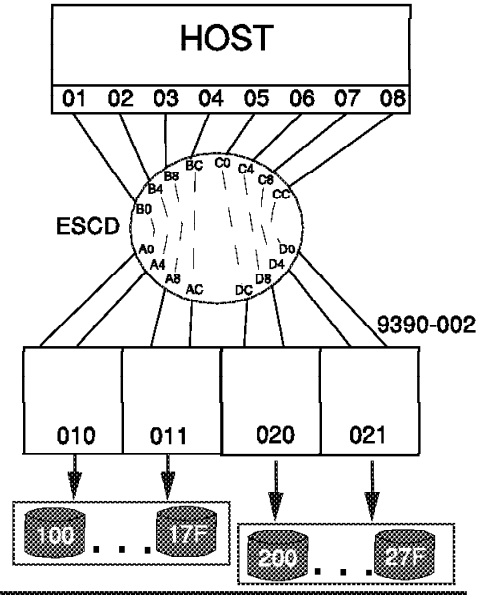
CNTLUNIT CUNUMBR=011,UNIT=3990
UNITADD=((00,128)),
PATH=(3,4)
LINK=(A8,AC)

IODEVICE CUNUMBR=(010,011),
UNIT=3390,UNITADD=00,
ADDRESS=(100,128)

CNTLUNIT CUNUMBR=020,UNIT=3990,
UNITADD=((00,128)),
PATH=(5,6)
LINK=(DC,D8)

CNTLUNIT CUNUMBR=021,UNIT=3990
UNITADD=((00,128)),
PATH=(7,8)
LINK=(D4,D0)

IODEVICE CUNUMBR=(020,021),
UNIT=3390,UNITADD=00,
ADDRESS=(200,128)



© IBM Corporation 1996

9390-002/ESCD IOCP Definition Example

The 9390-002 IOCP definitions for ESCD configurations are similar to the definitions for two 9390-001 configurations.

Physical Specifications



Type	Width		Depth		Height		Weight	
	mm	in.	mm	in.	mm	in.	kg	lb
9391-A30 empty rack	750	29.5	978	38.5	1580	62.2	341	752
9391-A30 full rack	750	29.5	978	38.5	1580	62.2	733	1616
9392-B33	212	8.3	780	30.7	175	6.9	24.5	54
9390	1155	45.5	815	32	1580	62.2	546	1204

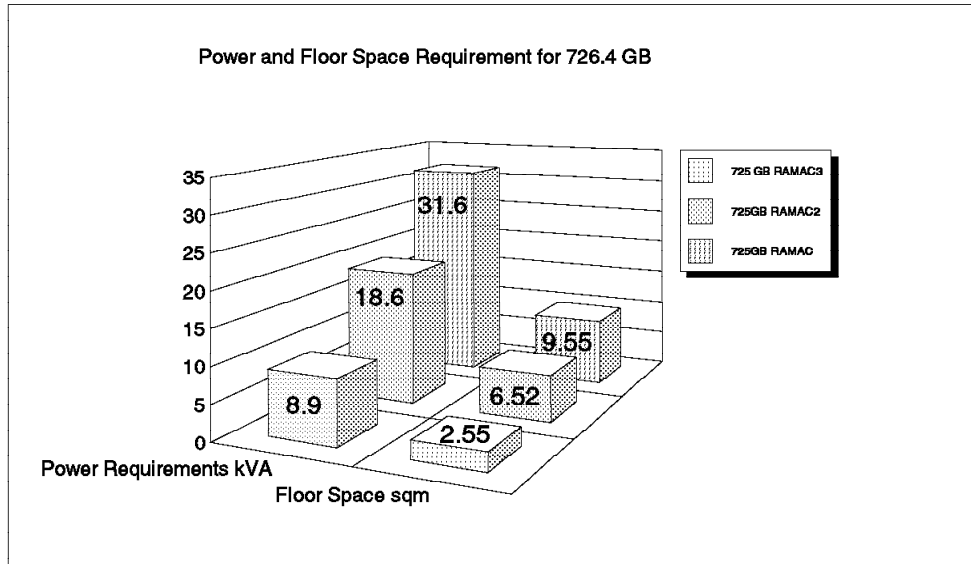


© IBM Corporation 1996

Physical Specifications

The foil shows the physical dimensions of the RAMAC 3 storage control, RAMAC 3 Storage Frame, and RAMAC 3 drawer.

Power and Floor Space Comparison

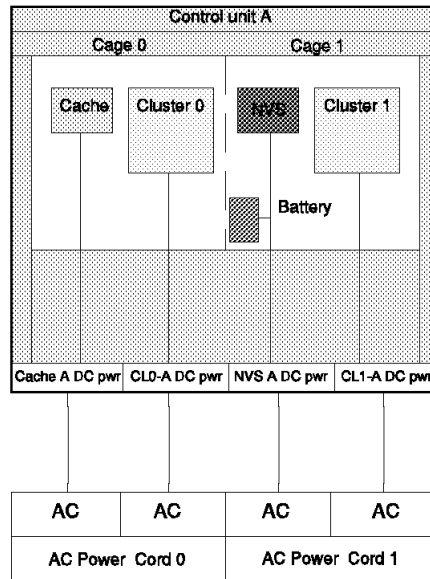


© IBM Corporation 1996

Power and Floor Space Comparison

The IBM RAMAC 3 Array Storage can significantly alleviate existing floor space constraints. The increased megabytes per square foot are made possible by the IBM Ultrastar 2XP 3.5 inch disk. Compared with the IBM RAMAC 2 Array DASD, the floor space is reduced by 50% for 726.4 GB of data. Also, the power and cooling requirements for the IBM RAMAC 3 Array Storage are reduced by 50%.

9390 Model 1 Power Boundaries



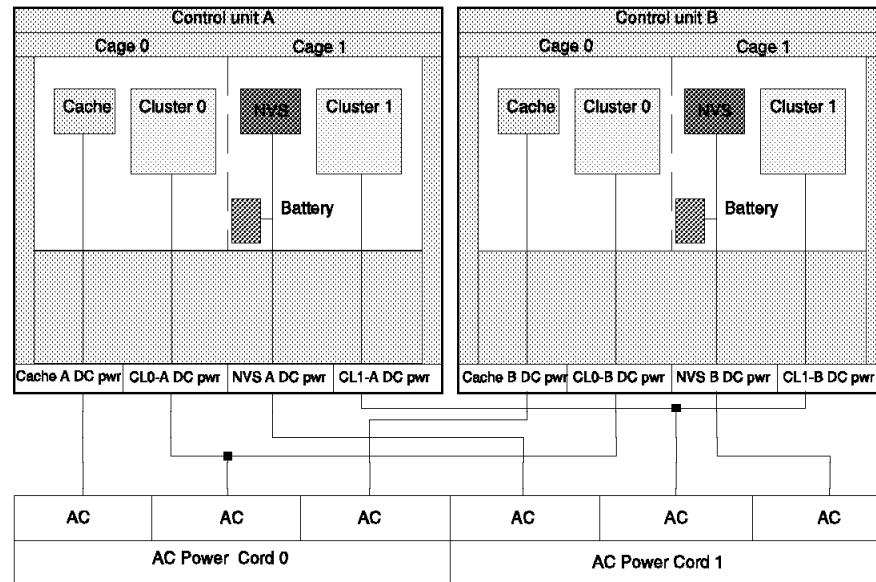
© IBM Corporation 1996

9390-001 Power Boundaries

The 9390-001 is a single control unit. The new redundant power subsystem is designed to enable nondisruptive upgrading to the 9390-002. The direct current (DC) power supply is repacked into new logic cards, which are installed in cage 0 and cage 1. Cage 0 holds the cache and cluster 0 DC power supply cards. Cage 1 holds the NVS and cluster 1 DC power supply cards.

The cooling system is repacked to fit in the new 9390. Power error handling is detected, recovered, and reported by the support facility (SF) microcode.

9390 Model 2 Power Boundaries



© IBM Corporation 1996

9390-002 Power Boundaries

The 9390-002 has a common emergency power-off switch for both control units. Each control unit has one subsystem power switch. To supply the second control unit with power, the alternating current (AC) power system has been expanded to six AC control boxes.

In case of failure of one of the redundant power systems, both control units are fully operable, but in some cases with reduced performance.

9390 Host Support



- ES/9000, S/390, or 3090 processors with ESCON channels
- ESCON support
 - Data transfer rate of up to 18 MB/s
 - 64 logical paths with 4 ports
 - 64 or 128 logical paths with 8 ports
 - 128 logical paths with 12 ports
 - 128 logical paths with 16 ports



© IBM Corporation 1996

9390 Host Support

The 9390-002 offers a total of 256 logical devices supported with 32 physical and 256 logical ESCON ports.

ES/9000, S/390, and 3090 processors with ESCON channels

The 9390 is supported by the 9672, 9021, 9121, 9291, and 3090 processors on ESCON channels only. The 9390 directly attaches to ESCON channels but does not directly attach to ESCON Extended Distance Feature (XDF) channels. To attach the 9390 to ESCON XDF channels, an ESCD or a 9036 must be positioned between the host processor and the 9390.

ESCON support

The maximum data transfer rate of the 9390 on an ESCON channel is 18 MB/s. The actual data transfer rate depends on the CPU ESCON channel data transfer rate.

When 64 logical paths are supported through two two-port ESCON cards, you have four ESCON channel ports.

To order use Feature code number 1400 or 2400.

When 128 logical paths are supported through four two-port ESCON cards, you have eight ESCON channel ports.

To order use Feature code number 1405 or 2405.

When 64 logical paths are supported through two four-port ESCON cards, you have eight ESCON channel ports.

To order use Feature code number 1410 or 2410.

When 128 logical paths are supported through two two-port and two four-port ESCON cards, you have 12 ESCON channel ports.

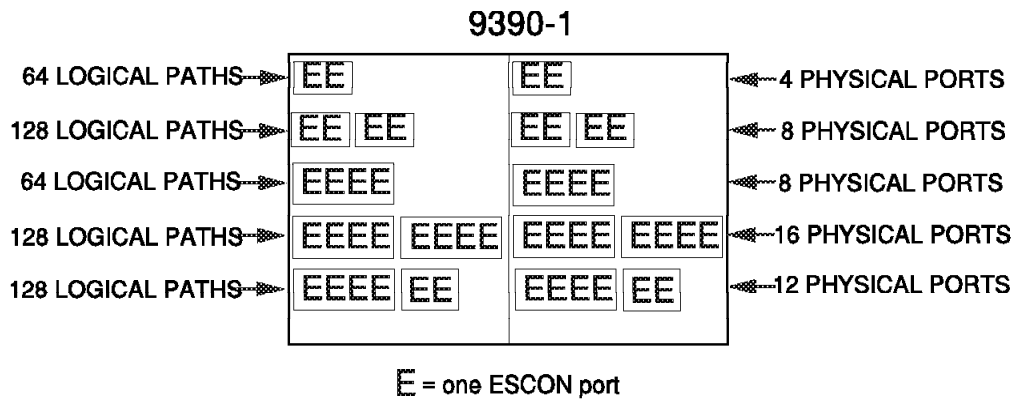
To order use Feature code number 1415 or 2415.

When 128 logical paths are supported through four four-port ESCON cards, you have 16 ESCON channel ports.

To order use Feature code number 1420 or 2420.

Only one feature code per 9390-001 and two per 9390-002 are installable. No combinations with other channel feature codes on a 9390-001 or 9390-002 are allowed.

9390-001 ESCON Channel Configuration



- 32 LOGICAL PATHS PER ESCON ADAPTER CARD
- 2 ESCON ADAPTER CARD TYPES - 2-PORT and 4-PORT
- CLUSTER SYMMETRY FOR ADAPTER CARD CONFIGURATION



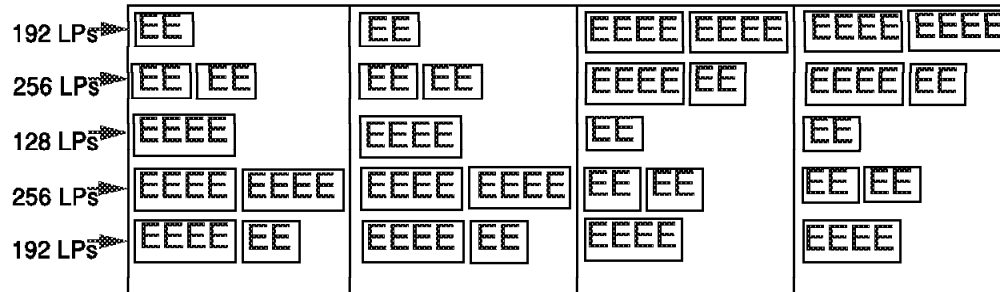
© IBM Corporation 1996

9390-001 ESCON Channel Configuration

The 9390-001 connects to the processor complex through ESCON channels, optical fiber cables, and optionally, ESCDs or 9036 extenders. The optical fiber cables connect to adapter cards located in each of the two clusters (four clusters for the 9390-002). The adapter cards have two or four ports; each port can connect a single optical fiber cable. Each ESCON adapter card supports up to 32 logical paths regardless of the number of physical ports on the card. Therefore when four cards are installed (two in each cluster), 128 logical paths are available for use. The distribution of logical paths across physical ports is determined by the IOCP definitions. Each cluster of the 9390-001 must have identical ESCON adapter card configurations—the foil shows the valid combinations.

9390-002 ESCON Channel Configuration...

9390-2



E = one ESCON port

▪ 9390-2 CONTROL UNIT ASYMMETRIC CONFIGURATIONS SUPPORTED



© IBM Corporation 1996

9390-002 ESCON Channel Configuration...

The 9390-002 comprises two storage controls in a single frame. Each storage control has two clusters that abide by the same rules established for the 9390-001 in terms of ESCON adapter card configuration. Each cluster within a single storage control must have identical ESCON adapter card configurations, but ESCON adapter card configurations can be different for each storage control. The foil shows how asymmetrical configurations are supported and illustrates the valid combinations of ESCON adapter cards. With a maximum of eight ESCON adapter cards installed, the 9390-002 supports up to 256 logical paths, distributed as two sets of 128 for each storage control. Logical path distribution is defined through IOCP parameter coding.

9390 Cache and NVS Feature Codes



Cache (MB)	9390-001	9390-002	NVS (MB)	9390-001	9390-002
256	1115	2115	32	1310	2310
512	1120	2120	64	1315	2315
768	1125	2125	128	1320	2320
1024	1130	2130			
1536	1135	2135			
2048	1140	2140			
3072	1145	2135			
4096	1150	2150			



© IBM Corporation 1996

9390 Cache and NVS Feature Codes

This table gives an overview of the available feature codes for the cache and NVS of the 9390.

3990-6 Support for RAMAC 3 Array



- Customer investment protection
- Faster DDC interface
- New LIC download and activation required for 128 device support



© IBM Corporation 1996

3990-6 Support for RAMAC 3 Array

Customer investment protection

To meet customer demand for investment protection, the 3990 Model 6 Storage Control has been enhanced to support 363.2 GB of RAMAC 3 Array capacity. The 3990 Model 6 Storage Control supports both parallel and ESCON channel attachment.

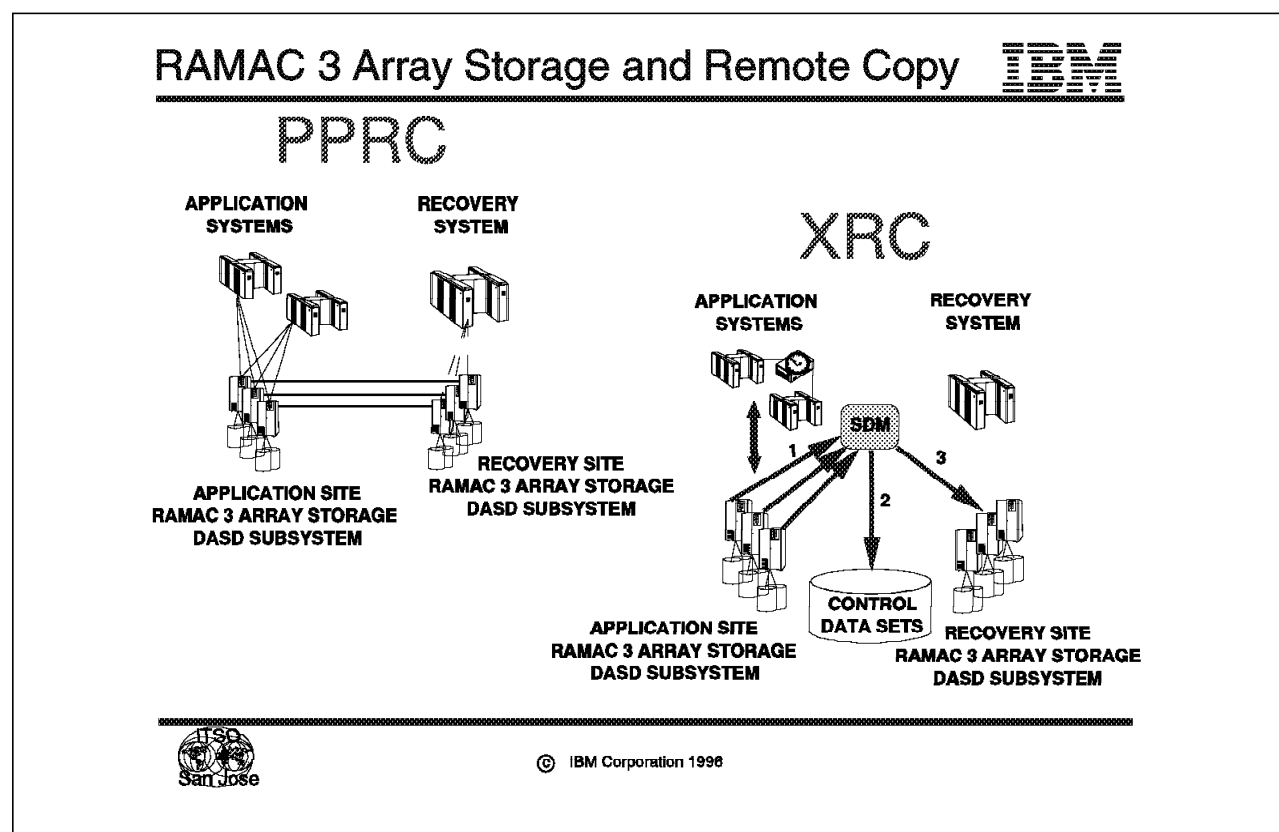
Faster DDC interface

When connected to the IBM RAMAC 3 Storage Frame, four new DDC cables are required for 3990-6 attachment. These cables must be ordered with the 9391-A30.

New LIC download and activation required for 128 device support

For RAMAC 3 drawer support, new Licensed Internal Code (LIC) must be ordered by the IBM customer engineer and loaded into the 3990 Model 6 Storage Control. This LIC supports 128-device addressing.

Chapter 8. RAMAC 3 Array Storage and Remote Copy

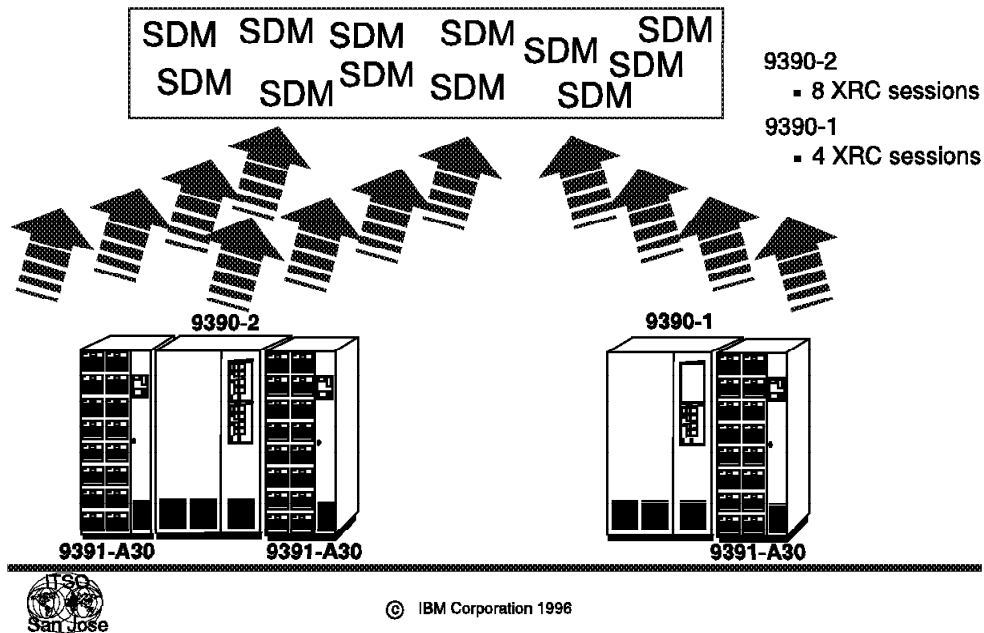


The 3990-6 supports two storage-based solutions that address workload movement, device migration, and disaster recovery—PPRC and XRC. This chapter describes the additional advantages that RAMAC 3 Array Storage offers for PPRC and XRC configurations.

9390 XRC Session Support



SUPPORTS COMPACT XRC-CU SESSION CONSOLIDATION



9390 XRC Session Support

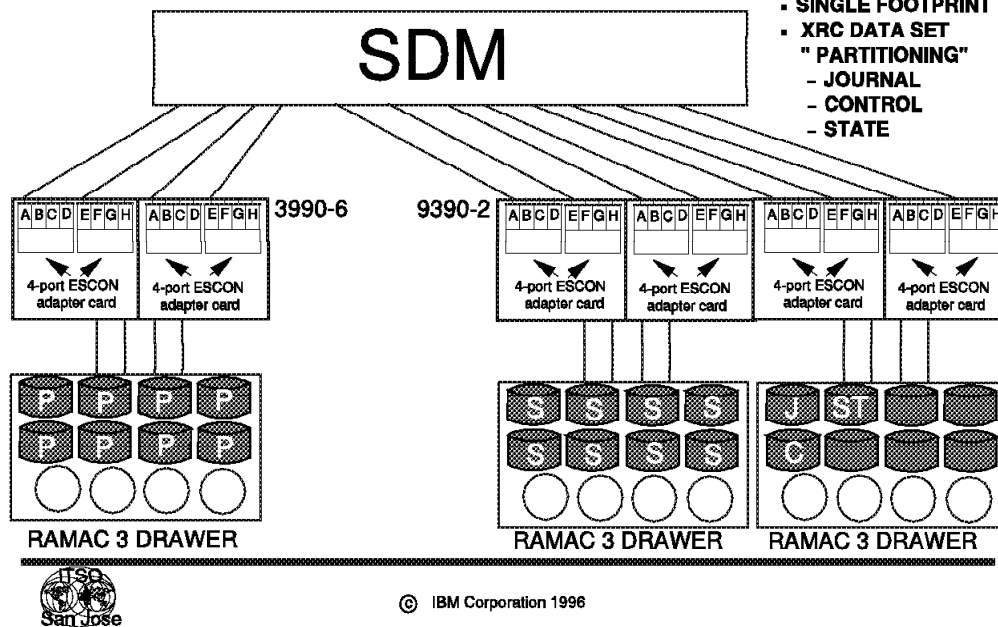
XRC provides a method of simplifying the management of real-time data shadowing for volumes that contain related data or where categories of data benefit from being grouped together. XRC does this by creating, managing, and monitoring sessions of data-shadowing activity. XRC primary and secondary volume pairs are added and deleted to XRC sessions, thus enabling XRC to simplify tasks such as monitoring, reporting, and recovery by reference to a single session identifier. Each XRC session requires a dedicated system data mover (SDM) to manage the remote copy logistics associated with the session. The logistics include real-time data shadowing, update sequence consistency processing, journaling, monitoring, and recovery. The XRC session concept enables SDM to manage the complexity of disaster recovery and data management by reference to a single name that relates many volumes with data dependencies. Support of the XRC session concept is vital and is delivered in the host software and the storage control.

The 3990-6 and the 9390-001 support up to four XRC sessions, and the 9390-002 supports up to eight XRC sessions. This support enables the most flexible and compact disaster recovery, workload movement, and data migration solution in the industry because large configurations can segregate corporate data into easily manageable XRC session "modules" for independent remote copy operations.

9390-2 at XRC Recovery Site



ENABLES OPTIMIZED XRC PERFORMANCE

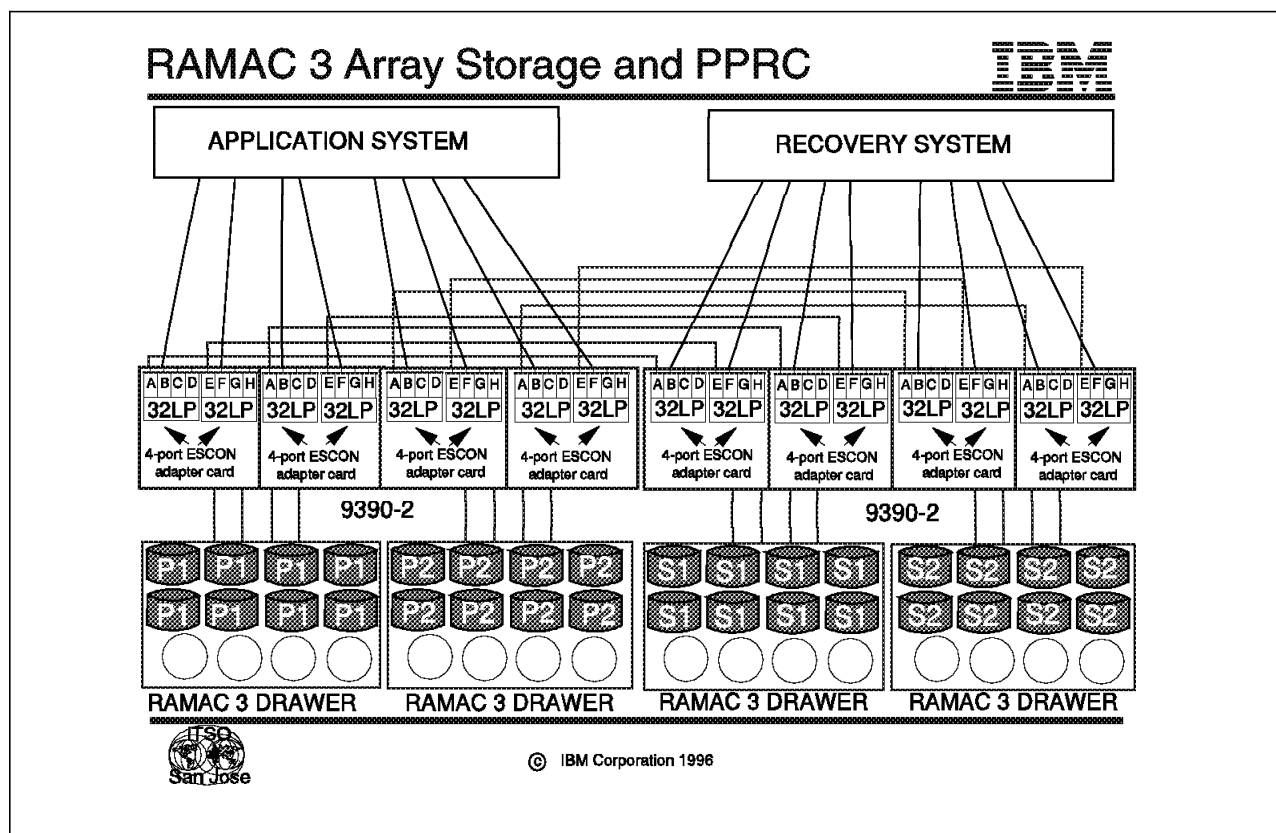


9390-002 at XRC Recovery Site

XRC application site storage controls must be either 3990-6, 9390-001, or 9390-002. XRC recovery site storage controls do not have to be 3990-6, 9390-001, or 9390-002—for example, the XRC recovery site storage control can be 3990-3, RAMAC Array Subsystem, or RAMAC Virtual Array. IBM recommends that the XRC recovery site storage control be a 3990-6, 9390-001, or 9390-002 because of their functional and performance attributes. However, the 9390-002 offers a unique capability that improves XRC performance and makes it a candidate for recovery site storage control.

XRC writes "batches" of updated records to a journal data set to enable subsequent secondary volume update in the event of disaster recovery and to rapidly secure a remote copy of the primary volume data. The batches of updated records are called *consistency groups*. The writing of consistency groups to the journal data set is a vital element of XRC processing and is on the "performance critical path." SDM will not release main storage space used to store record updates until the data is written to the secondary volumes. The data is not written to the secondary volumes until the consistency groups are stored on the journal data set. Therefore the performance of XRC journaling is important to the entire remote copy process. Because of the importance of journaling performance, we strongly recommend that the journal data set be "isolated" from the secondary volume update activity so that contention for storage control resources is minimized. This isolation effectively translates to a recommendation that the journal data sets associated with the XRC session be located on a different storage control from the secondary volumes. The 9390-002 is two storage controls in a single frame and therefore facilitates locating the

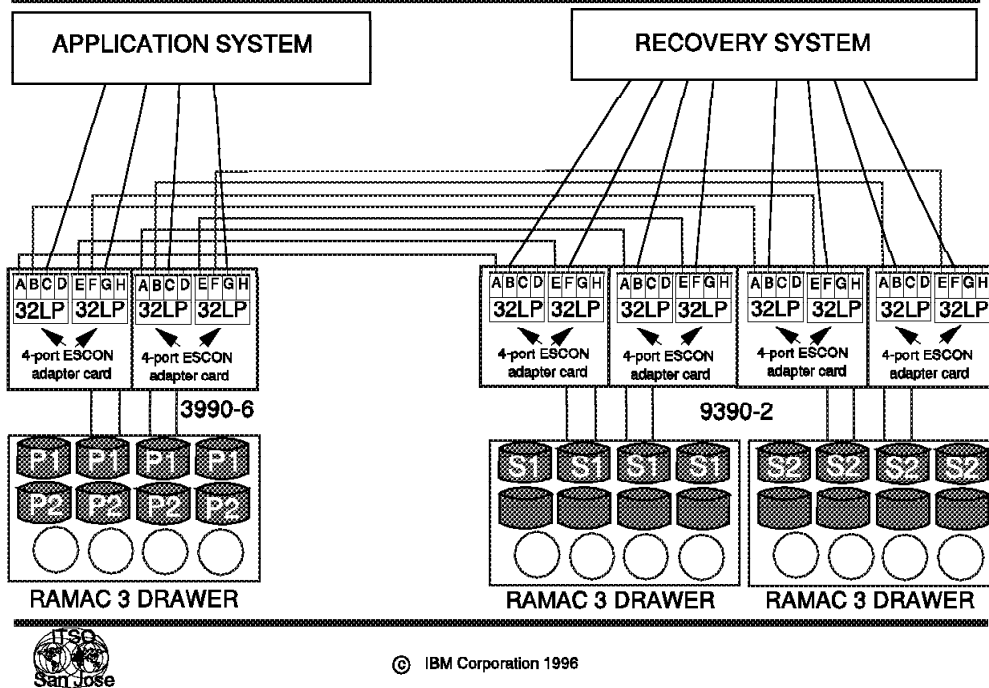
journal data sets in the same frame as the secondary volumes but behind a different storage control. The foil shows how the 9390-002 offers an economical, compact, and efficient XRC configuration by optimizing SDM journaling through a separation of journal data sets and secondary volumes.



RAMAC 3 Array Storage and PPRC

RAMAC 3 Array Storage supports up to 726.4 GB of data in a single frame and offers a level of high capacity compactness never before achievable in storage subsystem design. The compactness of a RAMAC 3 Array Storage subsystem delivers benefit to PPRC configurations that require real-time, synchronous, remote copy operations for disaster recovery, workload movement, and data migration. Before the RAMAC 3 Array drawer became available, a single 3990-6 could remotely copy a maximum of 180 GB of data stored on 64 volumes. Installations wanting to shadow more data than that by using PPRC had to use more 3990-6s. The 9390-002 enables a single frame footprint comprising 726.4 GB of data stored on 256 RAMAC 3 volumes to be remotely copied by using PPRC. Similarly, the 3990-6 attaching RAMAC 3 drawers enables a single-frame footprint comprising 363.2 GB of data stored on 128 RAMAC 3 volumes to be remotely copied by using PPRC. In addition, RAMAC 3 Array Storage and 3990-6 attaching RAMAC 3 drawers can exploit the advanced data management functions of P/DAS. The RAMAC 3 drawer is therefore the enabler for high-capacity, cost-effective, compact, remote copy configurations.

9390-002 at PPRC Recovery Site

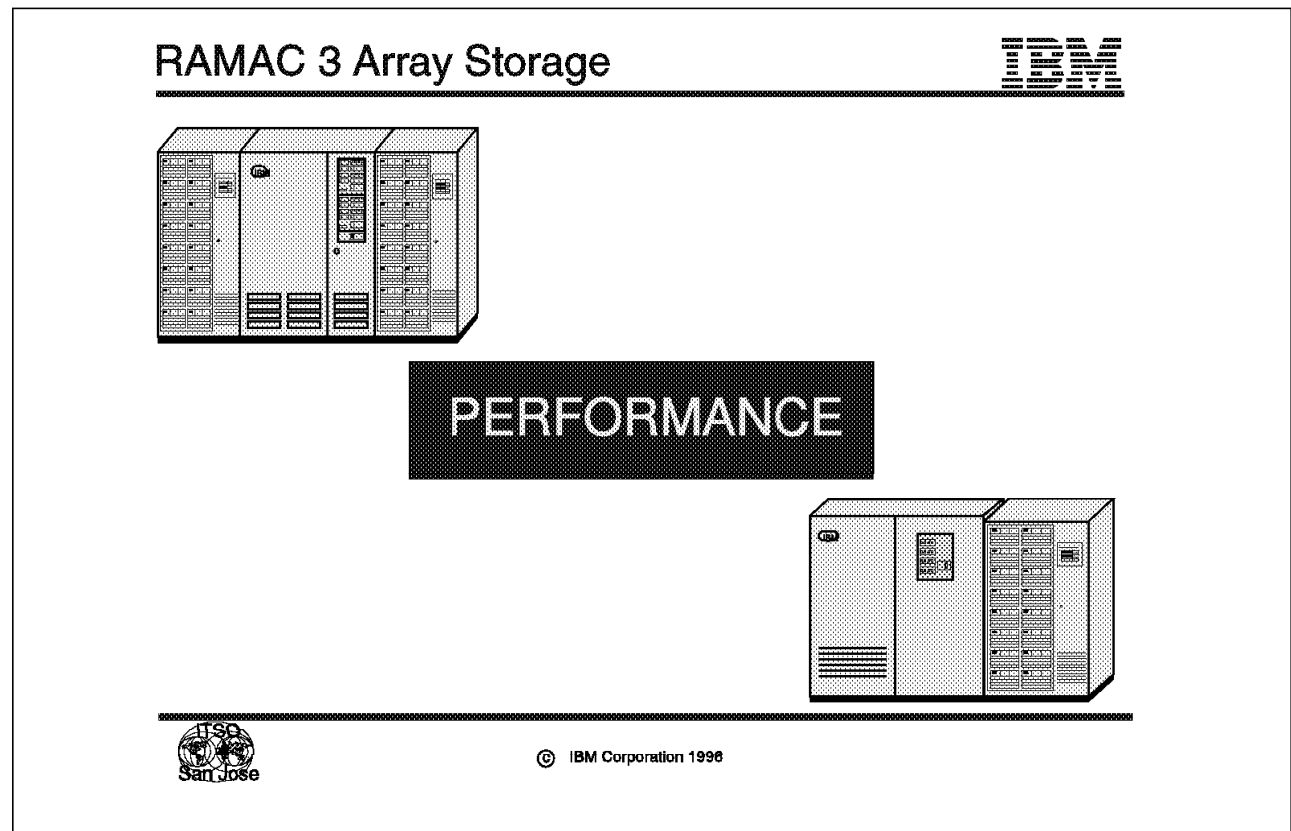


9390-002 at PPRC Recovery Site

Previously, PPRC configurations mandated the use of 3990-6 for both application and recovery site storage controls. With the arrival of the 9390 storage control, the options for PPRC configurations have been enhanced. PPRC performance is dependent on the write content of the application site storage control and the ability of the recovery site storage control to satisfy the shadowed write content. PPRC implements a synchronous remote copy operation so the write to the secondary volumes must execute as fast as possible to enable I/O completion status (device end—DE) to be returned to the application. Contention for recovery site storage control resources must therefore be removed or minimized so that acceptable PPRC performance can be sustained.

By using the 9390-002 as the recovery site storage control, secondary volumes can be "split" across control units. The split results in a reduction of the write content with which the control units have to deal. For high-write-content, application-site storage controls, the 9390-002 provides a way of balancing the shadowed workload across two storage control facilities at the recovery site, thus delivering optimal performance for the synchronous remote copy operations.

Chapter 9. Performance



RAMAC 3 Array Storage configurations support up to 726.4 GB of data in a single subsystem. Additionally, RAMAC 3 Array Storage subsystems support all of the functions currently available on the 3990-6 storage control. To enable all of the functionality and simultaneously support the massive-capacity footprint of a 726.4 GB subsystem, the RAMAC 3 Array Storage implements improvements in I/O processing and data transfer performance described previously. This chapter details some of the performance measurements undertaken with the RAMAC 3 Array Storage and highlights other performance aspects such as recommended cache and NVS configurations and use of IBM performance modeling tools.

Performance Modeling Tools



- DCAT
 - Migration support
 - *Response time*
 - *Throughput*
 - *Workload projection*
- CAA
 - Simulates hit ratios for cache and NVS sizes
 - Input to DCAT to evaluate cache and NVS effect on:
 - *Response time*
 - *Throughput*
 - *Workload projection*
- DASD Magic/2
 - Migration support
 - *Response time*
 - *Throughput*
- DSCAT
 - Data set level modeling

REMOVE THE RISK OF CHANGE
BY PREDICTING THE EFFECT
OF NEW TECHNOLOGY



© IBM Corporation 1996

Performance Tools

The effect of IBM RAMAC 3 Array Storage can be predicted through the use of IBM modeling tools. I/O response times, throughput, the effect of device migration, volume consolidation, and the effect of cache and NVS size increases can be modeled with the suite of IBM modeling tools.

DCAT

The Disk Cache Analysis Tool (DCAT) has been updated to support IBM RAMAC 3 Array Storage. DCAT enables IBM technical support personnel to model the impact that migrating to RAMAC 3 Array Storage has on response time and throughput. Workloads can be “stressed” by projecting a growth in I/O rates and modeling the resultant effect on subsystem performance.

CAA

An updated release of the Cache Analysis Aid (CAA) is available on MKTTOOLS. CAA uses generalized trace facility (GTF) input to simulate the cache hit ratios resulting from various cache sizes. The cache hit ratios are input to DCAT, which models the performance impact of changing cache size.

DASD Magic/2

DASD Magic/2 is a PC-based modeling tool that is used to predict the effect of device migration and cache size change. DASD Magic/2 can be used to model migration to RAMAC 3 Array Storage.

DSCAT

For data set and volume-specified performance analyses, use the Data Set Cache Analysis Tool (DSCAT), which you can find on MKTTOOLS.

Performance with Sequential Detect



RAMAC 3 Array Storage Sequential Read Performance

WORKLOAD	ELAPSED TIME RAMAC 2 ARRAY DASD	ELAPSED TIME RAMAC 3 ARRAY STORAGE
VSAM KSDS SEQ READ (NO SEQ DETECT)	22 MINUTES	N/A
VSAM KSDS SEQ READ (SEQ DETECT)	13 MINUTES	8 MINUTES

- 8 volumes active simultaneously across 4 drawers
- Each volume reads or writes 1500 cylinders
- 18 MB/s ESCON channel attachment
- 4 K block size



© IBM Corporation 1996

Performance with Sequential Detect

A critical area in many environments is sequential performance because it can be a significant portion of workload execution during batch processing. Demand for high throughput during these workloads continues to increase. The IBM RAMAC 3 Array Storage can provide significant improvements in this area over the IBM RAMAC Array DASD and IBM RAMAC 2 Array DASD.

With eight logical volumes of 3390-003 in a single drawer, maintaining and improving sequential throughput capability were key in the design of the IBM RAMAC 3 Array Storage. With the increased data rate to the IBM RAMAC 3 drawer, the minimum usable sequential throughput with all eight volumes in the disk array active is equal to or greater than that of the IBM RAMAC and RAMAC 2 Array DASD. When additional IBM RAMAC 3 Array Storage volumes are active, however, the overall throughput can be substantially higher than previously available. For a single 9390-001 or 3990-6 storage control, the maximum usable sequential throughput is in excess of 30 MB/s for read operations and 20 MB/s for write operations. This throughput is two to three times greater than previously available with IBM RAMAC Array DASD and RAMAC 2 Array DASD. For configurations involving the 9390-002, total subsystem throughput of 60 MB/s for reads and 40 MB/s for writes may be obtained. To achieve these results, improvements were made to the IBM RAMAC 3 Array Storage, and the 3990 Model 6 Storage Control LIC.

With sequential detect, workloads that previously had been unable to take advantage of the optimization for read sequential workloads can now do so. Applications that process data sequentially without setting the appropriate

software option for sequential access, such as VSAM key-sequenced data set (KSDS) processing and some OEM database applications, can now take advantage of sequential prestaging. In addition to improvements in elapsed time, the increased storage control to the RAMAC 3 drawer data rate also results in significantly lower utilization of the lower interface. This is significant on heavily loaded subsystems or on subsystems supporting greater than 181.6 GB of attached storage.

The foil indicates the improvements that may be possible with the RAMAC 3 Array Storage in sequential read processing environments.

Sequential Performance Improvements



RAMAC 3 Array Storage Sequential Write Performance

WORKLOAD	ELAPSED TIME RAMAC 2 ARRAY DASD	ELAPSED TIME RAMAC 3 ARRAY STORAGE
QSAM SEQ WRITE	16 MINUTES	9 MINUTES

- 8 volumes active simultaneously across 2 drawers
- Each volume reads or writes 1500 cylinders
- 18 MB/s ESCON channel attachment
- 27 K block size



© IBM Corporation 1996

Sequential Performance Improvements

The foil shows the improvements in sequential write performance that may be achieved with RAMAC 3 Array Storage for a QSAM workload.

RAMAC 3 Recommended Configurations



		Minimum		Performance	
Count of Volumes	Total GB	Total Cache Size (MB)	Total NVS Size (MB)	Total Cache Size (MB)	Total NVS Size (MB)
<=32	<= 90	256	32	512	32
32 - 64	90 - 180	512	32	1024 - 2048	64
64 - 128	180 - 360	1024	64	1024 - 4096	64
* 128 - 256	360 - 720	2048	128	2048 - 8192	256

Configurations for RAMAC 3 Array Storage attached to either 9390 or 3990-6 Storage Control

* denotes 9390-002 configuration



© IBM Corporation 1996

RAMAC 3 Recommended Configurations

This foil provides recommended standard and performance configurations for the IBM RAMAC 3 Array Storage when attached to either the 9390 or the 3990-6. The performance configurations may be applicable for workloads with high write content or where larger cache is necessary to improve cache hit ratios, such as may be found in some database applications.

When you evaluate any storage control configuration, we recommend that you analyze your workload to select the proper configuration for each environment.

Performance - Migration Consolidation



RAMAC 3 Array Storage Migration

CONFIGURATION	THROUGHPUT	RESPONSE TIME
2 X 3990-6 512 MB CACHE EACH 32 MB NVS EACH 180 GB RAMAC 2 EACH	275 I/Os/s EACH at 15 ms RT TOTAL 550 I/Os/s	14 ms RT EACH at 250 I/Os/s TOTAL 500 I/Os/s
1 X 9390-001 1 GB CACHE 64 MB NVS 360 GB RAMAC 3	580 I/Os/s at 15 ms RT	13 ms RT at 500 I/Os/s
1 X 9390-001 4 GB CACHE 128 MB NVS 360 GB RAMAC 3	770 I/Os/s at 10 ms RT	7 ms RT at 500 I/Os/s

DB2 Workload

- 3:1 R/W ratio
- 17 KB avg transfer
- 29% sequential stage
- 63% cache hit ratio at 1GB total cache
- 83% cache hit ratio at 4 GB total cache



© IBM Corporation 1996

Performance - Migration Consolidation

RAMAC 3 Array Storage configurations can support twice the capacity of previous 3990-6 configurations. Therefore, customers can migrate from multiple 3990-6 DASD subsystems to single RAMAC 3 Array Storage subsystems if appropriate capacity planning is done.

The foil shows the performance characteristics resulting from the consolidation of 360 GB attached to two 3990-6s to a single 9390-001. Each 3990-6 attaches 180 GB of RAMAC 2 capacity and sustains 275 I/Os/s with an average response time of 15 ms. After migrating all of the RAMAC 2 drawer capacity attached to the two 3990-6s to RAMAC 3 drawers attached to the 9390-001, the average response time is maintained with an increase in throughput. The faster lower-interface data transfer capability of RAMAC 3 Array Storage, and the subsequent effect on larger cache and NVS effectiveness, has eliminated the constraint of high DB2 sequential prestaging and low cache hit ratios that would normally have prevailed for this workload.

Chapter 10. Software

RAMAC 3 Array Storage

IBM
RACAL
DATA
PROCESSING
CORPORATION
SAN JOSE
CALIFORNIA
95128-1600
U.S.A.



© IBM Corporation 1996

This chapter describes the software support required to implement RAMAC 3 Array Storage configurations. The software changes focus on the support for 128 and 256 device addresses that is now provided by 3990-6, 9390-001, and 9390-002 storage controls. Enhancements to software utilities and storage subsystem monitoring functions are also covered. Check the preventative service planning (PSP) records for up-to-date APARs and PTFs.

RAMAC 3 Array Storage Software Support



MVS

MVS/ESA SP V5 R2.2
MVS/ESA SP V5 R2.0
MVS/ESA SP V5 R1.0
MVS/ESA SP V4 R3.0
MVS/ESA SP V4 R2.2
MVS/ESA SP V4 R2.0
MVS/SP V3 R1.3

DFSMS/DFP

DFSMS/MVS V1 R3.0
DFSMS/MVS V1 R2.0
DFSMS/MVS V1 R1.0
MVS/DFP V3 R3.2
MVS/DFP V3 R3.1
MVS/DFP V3 R3.0
MVS/DFP V3 R1.1

VM/ESA

VM/ESA V2 R1.0
VM/ESA V1 R2.2

VSE/ESA

VSE/ESA V2 R1	
VSE/ESA V1 R4	(1)
VSE/ESA V1 R3	(1)

(1) Support provided only for the 9390 and 3990-6 in basic mode



© IBM Corporation 1996

RAMAC 3 Array Storage Software Support

This foil shows the supported SCPs and DFSMS and DFP versions.

DATE	DESCRIPTION	AMOUNT	BALANCE
1980	1/1		100.00
1980	2/1	50.00	150.00
1980	3/1	25.00	175.00
1980	4/1	75.00	250.00
1980	5/1	100.00	350.00
1980	6/1	150.00	500.00
1980	7/1	200.00	700.00
1980	8/1	250.00	950.00
1980	9/1	300.00	1250.00
1980	10/1	350.00	1600.00
1980	11/1	400.00	2000.00
1980	12/1	450.00	2450.00
1980	12/31		2450.00

AOM SUPPORT FOR 256 ADDRESSES		APARs	
		OW22821	OW22823
		↓	↓
		PTFs	
DFSMS 1.1.0	UW32444		
DFSMS 1.2.0	UW32442		
DFSMS 1.3.0	UW32443		
DFP 3.3.0			UW32445



© IBM Corporation 1996

[illegible]

	APARs	
CUIR	OW23139	OW23141
	PTFs	
DFSMS 1.1.0	UW32938	
DFSMS 1.2.0	UW32936	
DFSMS 1.3.0	UW32937	
DFP 3.3.0		UW33079



© IBM Corporation 1996

142 IBM RAMAC 3 Array Storage

RAMAC 3 Software - APARs and PTFs...



APARs		
DFDSS	OW22109	PN87880
	PTFs	
DFSMSdss 1.1.0	UW32558	
DFSMSdss 1.2.0	UW32556	
DFSMSdss 1.3.0	UW32557	
DFDSS Ver 2.5		UN95927

ICKDSF R16 (incl 3380 TF)	APAR - PN85631
	PTF
MVS	UN92119
VM	UN92118
VSE	UN92122

ERE3.5	APAR	PTF
MVS	IR33151	UR46466
VM	VM60764	UV59774
VSE	DY44291	UD50203



© IBM Corporation 1996

RAMAC 3 Software - APARs and PTFs...

The APARs and PTFs listed enable DFDSS, DFSMSdss, EREP, and ICKDSF to recognize the physical characteristics of the new RAMAC 3 device even though it emulates a 3390-3 volume.

RAMAC 3 Software - APARs and PTFs...



APARs		
DEVSERV DFSMS	OW23106	OW23152
	↓	↓
	PTFs	
DFSMS 1.1.0	UW32926	
DFSMS 1.2.0	UW32924	
DFSMS 1.3.0	UW32925	
DFP 3.3.0		UW32928
DFP 3.3.2		UW32929
DFP 3.1.1		UW32927

APARs		
DEVSERV (RTYPE)	OW22489	OW22865
	↓	↓
	PTFs	
DFSMS 1.1.0	UW32980	
DFSMS 1.2.0	UW32978	
DFSMS 1.3.0	UW32979	
DFP 3.3.0		UW33049
DFP 3.3.2		UW33049

APARs		
DEVSERV DFSMS	OW22730	OW22731
	↓	↓
	PTFs	
DFSMS 1.1.0	UW32461	
DFSMS 1.2.0	UW32459	
DFSMS 1.3.0	UW32460	
DFP 3.3.0		UW32457
DFP 3.3.2		UW32458
DFP 3.1.1		UW32456

DEVSERV (SIM)	APAR - VM60783
	PTF
VM/ESA 1.2.2	UM28020
VM/ESA 2.1.0	UM28019



© IBM Corporation 1996

RAMAC 3 Software - APARs and PTFs...

The MVS DEVSERV command reports the physical device type for RAMAC 3 Array drawer configurations (the DEVSERV report is detailed in Chapter 11). Support for DEVSERV reporting of RAMAC 3 drawer physical device type is included in the APARs and PTFs listed.

RAMAC 3 Array Storage and RAMAC 3 Array DASD configurations communicate their device characteristics to the host to enable DFSMS to allocate data sets according to availability and performance specifications. The APARs and PTFs that allow DFSMS to recognize the new RAMAC 3 Array drawer as a high-performance RAID-5 device are listed.

Chapter 11. Operational Considerations

RAMAC 3 Array Storage



Operational Considerations



© IBM Corporation 1996

Operations staff must have adequate reporting and monitoring assistance when installations implement new technologies. For RAMAC 3 Array Storage, a simple change to the DEVSERV command output helps operators identify the new RAMAC 3 Array drawer. The details are covered in this chapter.

RAMAC 3 DEVSERV Format

0000 000000 0000 0000
0000 000000 0000 0000
0000 000000 0000 0000
0000 000000 0000 0000
0000 000000 0000 0000
0000 000000 0000 0000
0000 000000 0000 0000
0000 000000 0000 0000

**DEVSERV reports RAMAC 3 real device type
RTYPE 9392-3**

DS P,F41

IEE459I 09.15.55 DEVSERV PATHS 106

UNIT DTYPE M CNT VOLSER CHPID=PATH STATUS

RTYPE SSID CFW TC DFW PIN DC-STATE CCA DDC ALT CU-TYPE

0F41,33903 ,0,000, 220RES, B1=+ C1=+ B7=+ C7=+

9392-3 0057 Y YY. NY. N SIMPLEX 00 40 3990-6

***** SYMBOL DEFINITIONS *****

0 = ONLINE

+ = PATH AVAILABLE



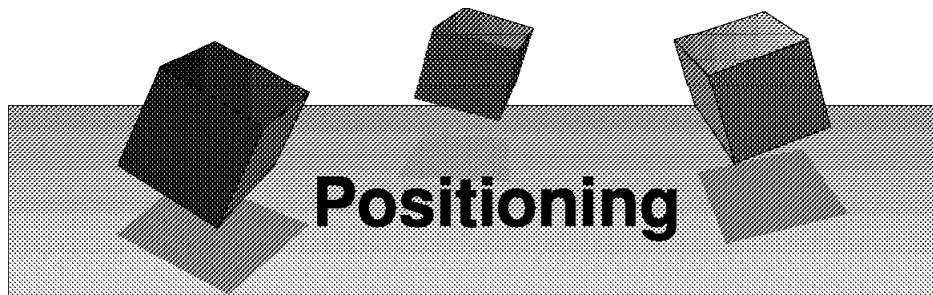
© IBM Corporation 1996

RAMAC 3 DEVSERV Format

The DEVSERV MVS console command has been enhanced to display the real device type for RAMAC 3 drawers. RAMAC, RAMAC 2, and RAMAC 3 drawers emulate 3390-3 device geometry—this appears in the DEVSERV report under the DTYPE column. The real device type for RAMAC 3 drawers is reported as 9392-3 under the RTYPE column. The APARs and PTFs that should be applied to deliver this reporting enhancement are listed on the foil.

Chapter 12. Positioning

RAMAC 3 Array Storage



© IBM Corporation 1996

The RAMAC Array Family has been expanded by no fewer than four new products: RAMAC 3 Array Storage, the RAMAC Virtual Array, the RAMAC Scalable Array, and the RAMAC Electronic Array. Each of these products provides a level of cost-effectiveness balanced with functionality that can be mapped against all of the storage subsystem requirements in most installations today. In this chapter we describe some of the RAMAC Array Family product characteristics that you can use to evaluate the “fit of the products” to your various information technology strategies and requirements.

RAMAC 3 Array Storage Positioning



- RAMAC 3 Array Storage is designed for:
 - Highest data availability in the industry, 7 days 24 hours
 - Fault-tolerant RAID-5 implementation using third-generation RAMAC drawer technology
 - Best disaster recovery with PPRC or XRC
 - High performance in sequential and random data access
 - Up to 726.4 GB fault tolerant data storage



© IBM Corporation 1996

RAMAC 3 Array Storage Positioning

The IBM RAMAC Array Family is the industry leader in RAID-5 technologies with more than 11,000 RAMAC Array Family subsystems installed worldwide. IBM continues its leadership in the industry by introducing new storage products developed through the exploitation of the richest technology base available.

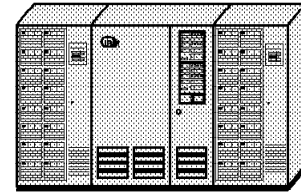
RAMAC 3 Array Storage is designed for:

The third-generation RAID-5 implementation with redundant parts and drawer sparing capability is the high-end storage product for the System/390 environment. PPRC and XRC provide ideal data protection for disaster recovery. P/DAS provides additional operational flexibility for data migration. Sequential data access has been improved compared with the IBM RAMAC 2 Array DASD. This improvement will optimize all VSAM sequential processing. The IBM RAMAC 3 Array Storage can be incrementally upgraded from 45.4 GB to 726.4 GB in steps of 22.7 GB.

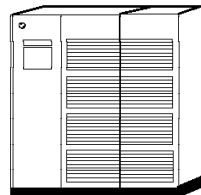
Positioning of New RAMAC Array Products



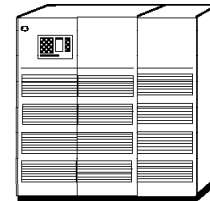
- **RAMAC 3 Array Storage**
 - Highest availability
 - Remote copy
 - High performance
- **RAMAC Virtual Array**
 - Excellent storage management
 - SnapShot
 - Low cost per GB
- **RAMAC Scalable Array**
 - High throughput
 - High availability
 - Up to 16 host data paths



RAMAC 3 Array Storage
726.4 GB



RAMAC Virtual Array
726 GB



RAMAC Scalable Array
1.4 TB



© IBM Corporation 1996

Positioning of New RAMAC Array Products

The RAMAC Array Family offers a variety of storage features.

RAMAC 3 Array Storage

The RAMAC 3 Array Storage provides the highest levels of storage subsystem performance at the highest levels of data availability. The RAMAC 3 Array Storage offers industry-leading reliability and the most complete portfolio of solutions that range from flexible configuration design and maintenance, through workload movement, device migration, and disaster recovery.

RAMAC Virtual Array

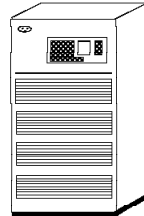
The RAMAC Virtual Array, with flexibility provided through its log structure file architecture and self-tuning performance, is the right solution for a broad range of your tasks. SnapShot with software support offers fast data replication with no need for more data space until the data is written. The RAMAC Virtual Array takes the smallest amount of floor space per gigabyte and has broad flexibility in virtual device support, with up to 256 addresses. It supports 3380 and 3390 device types with flexible volume sizes. Its RAID-6 architecture with dual-disk-drive redundancy helps to maintain high availability. The unique operational flexibility and low cost of ownership are the key advantages of this data storage subsystem.

RAMAC Scalable Array

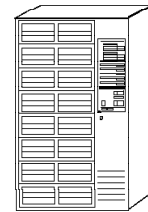
The fault-tolerant RAID-5 RAMAC Scalable Array is designed for high throughput and excellent performance at high I/O rates. With 16 data transfer paths and up to 4 GB of mirrored cache, you can have up to 1.4 TB of data storage. Disaster recovery is supported through XRC.

Positioning of New RAMAC Array Products...

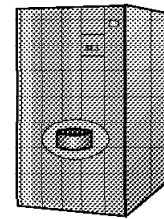
- **RAMAC 2 Subsystem**
 - Highest-availability RAID-5
 - Entry storage
 - High performance
- **RAMAC Electronic Array**
 - Highest performance
 - Up to 16 host data paths
 - Mirrored nonvolatile storage
- **S/390 Internal Disk**
 - High performance
 - High-availability RAID-1
 - Integrated disk server with lowest cost per GB



RAMAC Electronic Array
4 GB



RAMAC 2 Subsystem
180 GB



S/390 Internal Disk
288 GB



© IBM Corporation 1996

Positioning of RAMAC Family...

RAMAC 2 Array Subsystem

The RAMAC 2 Array Subsystem is available as a Model 1, 2, or 3. It is the RAMAC entry system for high-performance RAID-5 data storage. It delivers investment protection for RAMAC Storage with mixed RAMAC and RAMAC 2 drawers for users with low data rate growth.

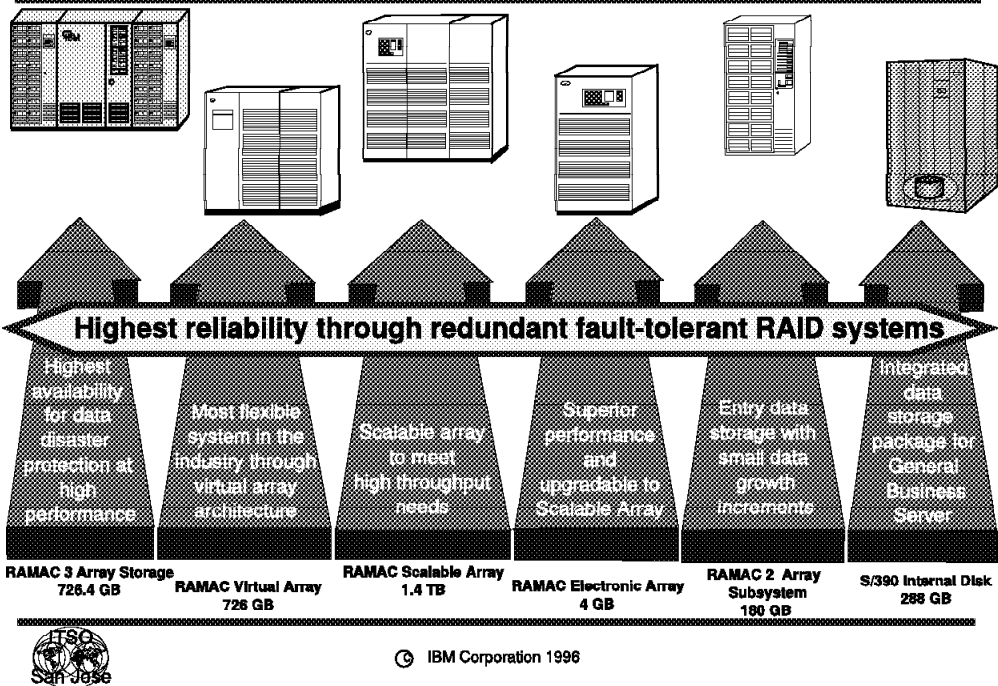
RAMAC Electronic Array

The RAMAC Electronic Array delivers superior performance by providing data response times at cache speed. It has up to 4 GB of memory, up to 16 data paths and two 100 MB data buses. The RAMAC Electronic Array can be upgraded to the RAMAC Scalable Array.

S/390 Internal Disk

The S/390 Internal Disk is the most cost-efficient and floor-space-efficient solution for storage. It has exceptionally low environmental requirements. The RAID-1 S/390 Internal Disk provides unique storage packaging for the General Business Server (GBS) and is the newest member of IBM's storage solutions family. The drawer capacity can be incremented from 18 GB to 288 GB in steps of 18 GB. Cache support is provided through processor memory and varies from 32 MB to 1 GB. The S/390 Internal Disk offers virtual disks in 3380 track format and 3390 track format. As a solution offering option, an OS/390, VM, or VSE SCP can be preloaded onto the S/390 Internal Disk.

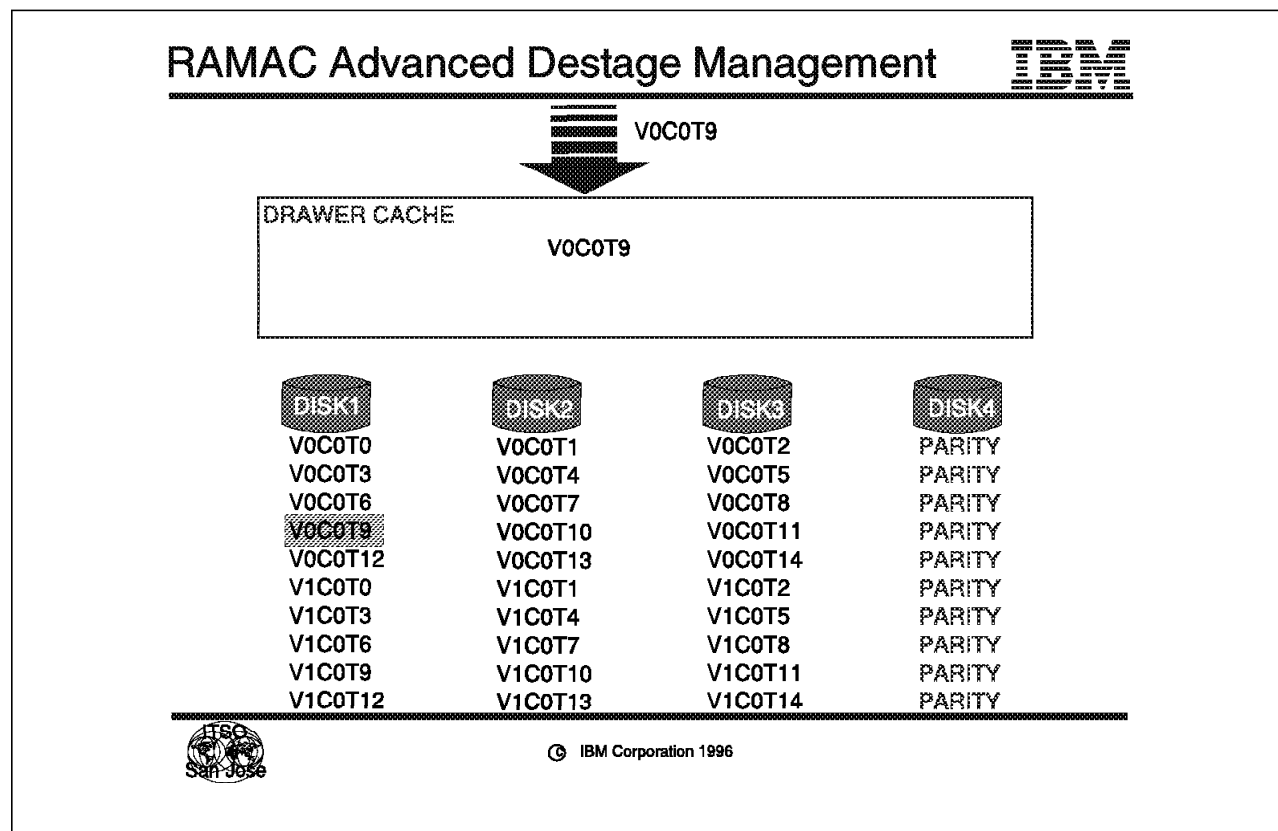
IBM RAID Data Storage Offerings for S/390



IBM RAID Data Storage Offerings for S/390

IBM provides the broadest range of S/390 storage solutions in the industry. IBM's solutions can meet any of your requirements—and all deliver high-availability and high performance.

Appendix A. RAMAC Advanced Destage Management



RAMAC Advanced Destage Management

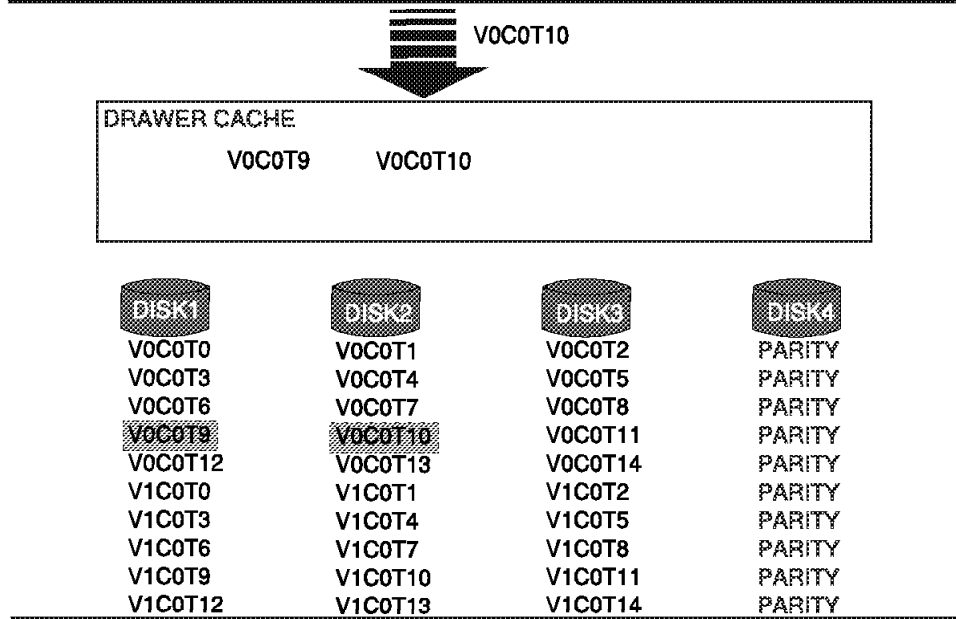
The RAMAC, RAMAC 2, and RAMAC 3 drawers implement an advanced destage management function that optimizes RAID-5 operations by avoiding the write penalty associated with parity computations for sequential workloads. The foil shows the first track of a sequential operation being updated in drawer cache—the track is on logical volume 0, cylinder 0, track 9 (V0C0T9). As the track is written to drawer cache, the expectation is that the old version of the track is staged to drawer cache along with the old version of parity associated with that track, and the exclusive OR (XOR) operation is executed against the old and new data and the old parity to generate the new parity. After the new parity is calculated, the new version of the track and the new version of parity are written to their appropriate disk drives in the array. The traditional RAID-5 operation described above involves two reads and two writes (read the old data and the old parity, then write the new data and the new parity)—to preserve redundancy in the array.

RAMAC drawers are not dictated to by conventional RAID-5 techniques. When sequential intent is indicated in the channel program, the RAMAC drawer “delays” the parity processing activity, thus seemingly delaying the staging of old data and old parity. When the RAMAC drawer receives the first track into drawer cache, it waits for the adjacent tracks to arrive “soon after.” This is a

reasonable approach, considering that sequential intent has been signaled and the next track in sequence must be written to drawer cache. The foil highlights track V0C0T9 in drawer cache and its intended location on disk 1 when it is eventually destaged—but not yet! The next foil describes how the drawer reacts to the next track of the sequential data stream.

RAMAC Advanced Destage Management...

IBM
San Jose

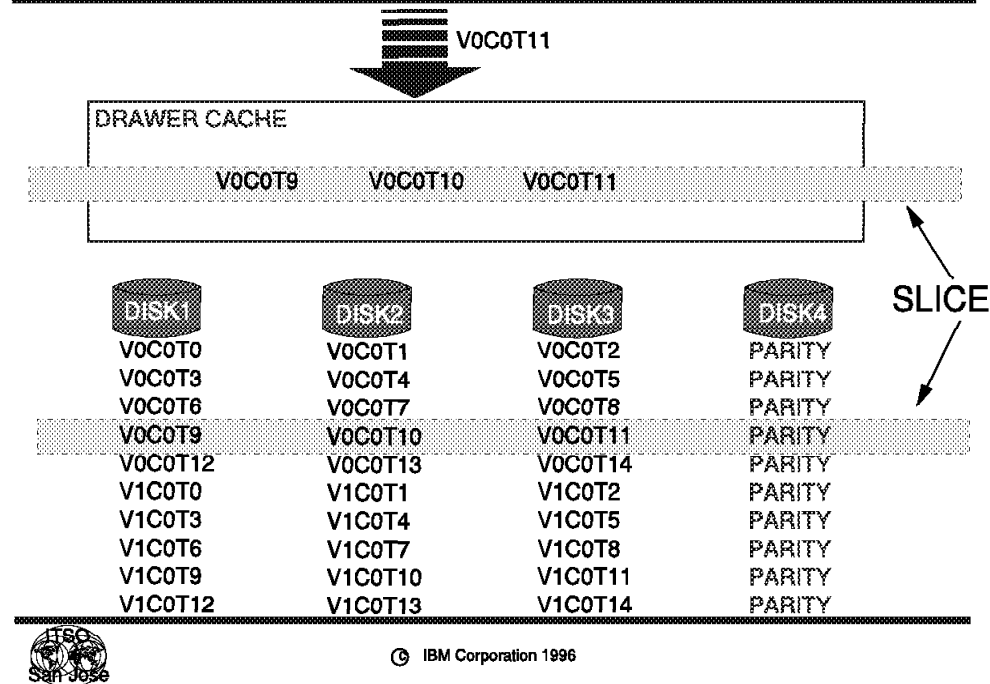


© IBM Corporation 1996

RAMAC Advanced Destage Management...

When the second track is written to the RAMAC drawer, the first track is already there but not yet written to disk. The second track is logical volume 0, cylinder 0, track 10 (V0C0T10). The foil shows both tracks now located in drawer cache with their eventual disk drive locations highlighted—V0C0T9 destined for disk 1 and V0C0T10 destined for disk 2. Refer now to the next foil for an explanation of why the parity processing and associated overhead of conventional RAID-5 can be avoided.

RAMAC Advanced Destage Management...

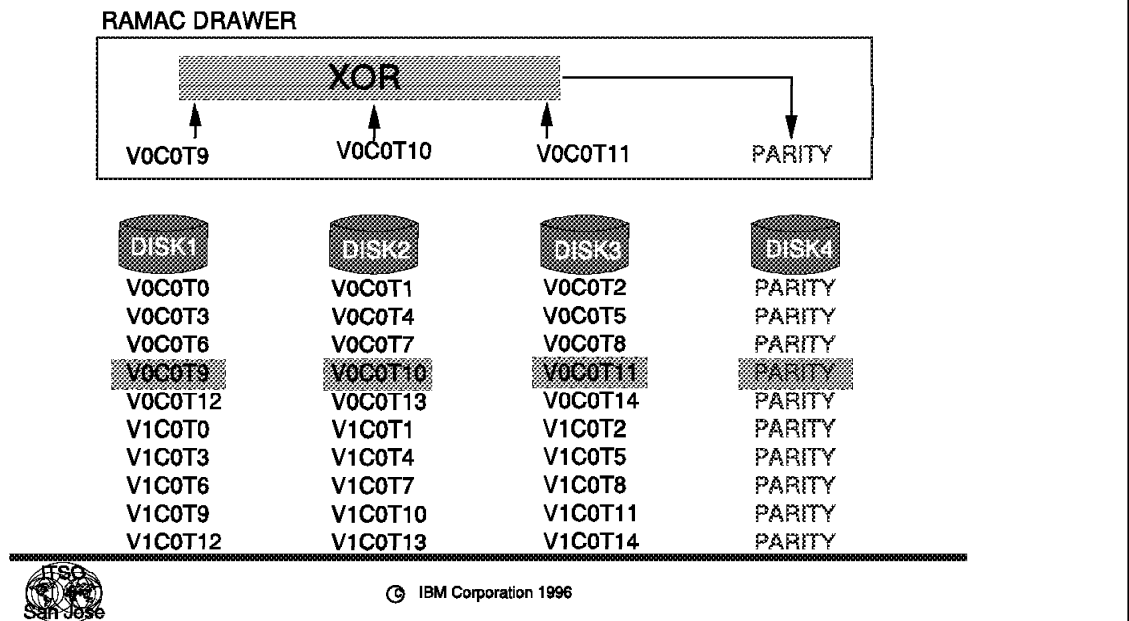


RAMAC Advanced Destage Management...

The foil shows the third track written to the RAMAC drawer in sequence—V0C0T11. When this track has been written to the drawer cache, the drawer has all of the information necessary to calculate the parity associated with the three data tracks. This is because a single XOR of three data tracks in the same slice produces the parity information for that slice. A slice is comprised of three adjacent data tracks—tracks 0, 1, and 2; tracks 3, 4, and 5; tracks 6, 7, and 8; tracks 9, 10, and 11; and tracks 12, 13, and 14—for the same logical cylinder will always occupy the same locations on the disk drives and will always comprise a slice. The RAMAC drawer knows in advance that sequential intent means that adjacent tracks will be written and will “suppress” the conventional parity processing actions needed for data written to the drawer. The next foil illustrates how parity computations produce no write penalty for RAMAC drawer sequential processing.

RAMAC Advanced Destage Management... **IBM**

WRITE PENALTY AVOIDANCE FOR SEQUENTIAL I/O

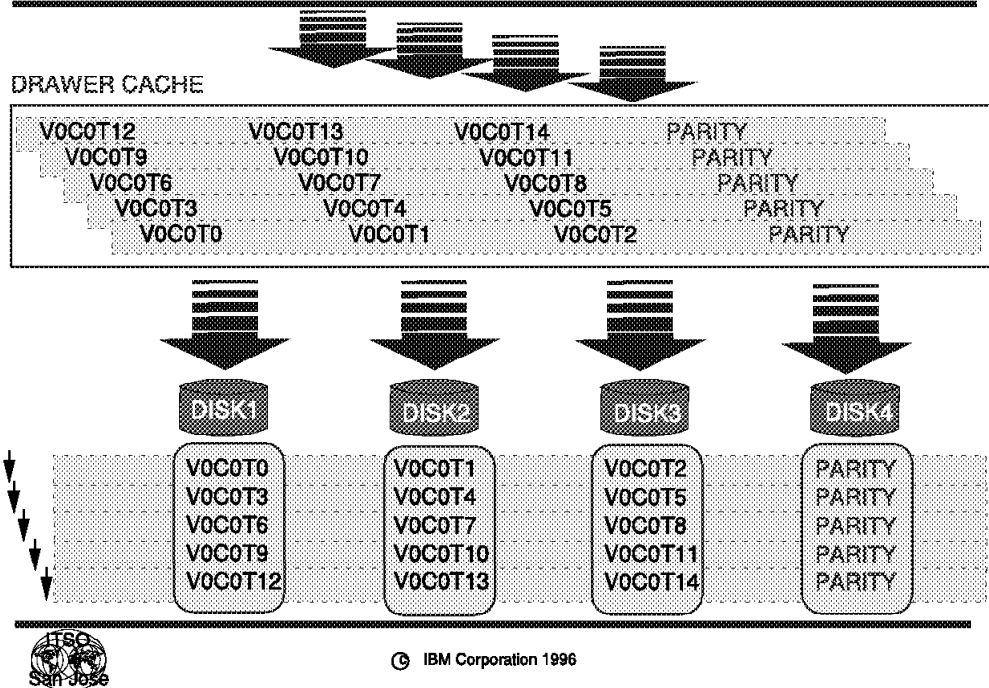


RAMAC Advanced Destage Management...

The RAMAC drawer has all three adjacent tracks in the same slice and now initiates the XOR operation. The RAMAC drawer XOR function is delivered by fast hardware engines that accept three binary data streams as input (three adjacent data tracks) and produces a single binary data stream as output (the new parity track associated with the three data tracks). At this point, the RAMAC drawer has computed the parity information that will protect against data loss should one of the disk drives fail—this was accomplished without incurring the read of old data and old parity into the drawer cache. Thus parity has been calculated “free of charge” in the RAMAC drawer.

The foil highlights the RAMAC drawer’s generation of the parity information using its XOR hardware engine and shows that the RAMAC drawer contains the three data tracks and the parity track in drawer cache. Also highlighted are the four locations on the disk drives of where these tracks will be written—notice that the parity track is written on disk 4. The four tracks have not yet been written to their disk drive locations, but the drawer cache is nonvolatile, so the contents are protected in the event of power loss. The next phase of the drawer’s advanced destage management technique is to optimize the writing to disk of the data and parity information and thus close the loop on enhanced RAID-5 execution.

RAMAC Advanced Destage Management...



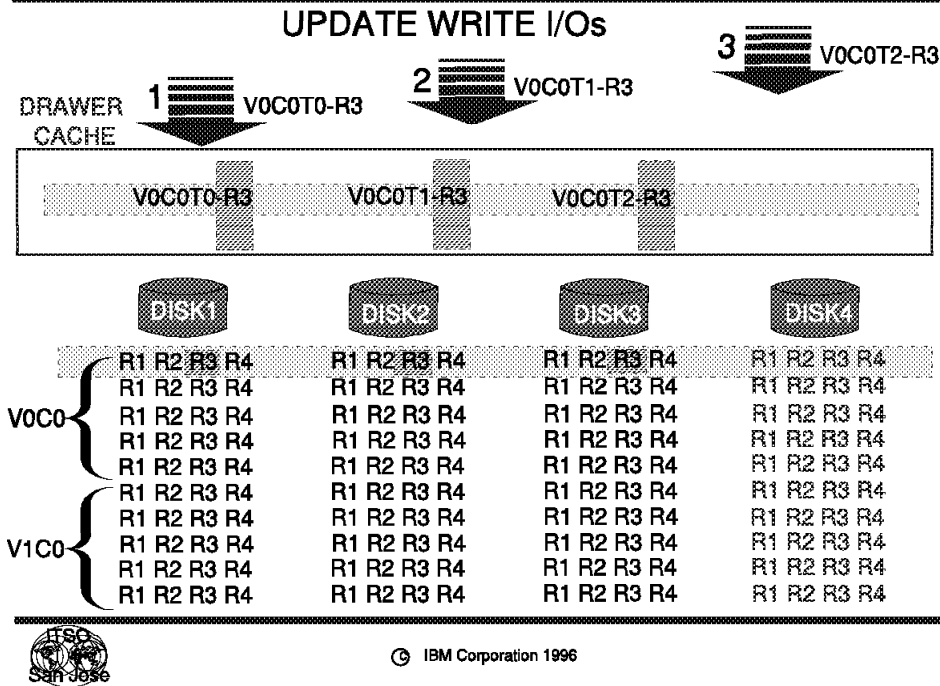
RAMAC Advanced Destage Management...

The previous foils demonstrate that the RAMAC, RAMAC 2, and RAMAC 3 drawers provide RAID-5 function without RAID-5 penalty—thus we describe the RAMAC drawer as an enhanced RAID-5 device. Data updated in nonvolatile semiconductor storage must ultimately be written to disk—in a RAID-5 implementation, this entails writing to multiple disks. The RAMAC Array Family drawer design provides yet another technique designed to optimize performance when destaging data.

RAMAC 2 and RAMAC 3 drawer advanced destaging techniques overcome the problem of inefficient disk use caused by random destaging. Rather than initiate a destage from drawer cache to the disk drives on the basis of frequency of use of the data, the RAMAC 2 and RAMAC 3 drawers set a track threshold for each disk. When the threshold is reached, it groups the data by disk drive location and writes the data out in such a way as to avoid repeated arm positioning.

The foil depicts the destage operation. Notice that the tracks are "sorted" into ascending SCSI block numbers for each disk to avoid disk arm "thrashing" effects.

RAMAC Advanced Destage Management



RAMAC Advanced Destage Management...

The RAMAC drawers take advantage of a highly parallel design that provides hardware-assisted enhanced RAID-5 function while exploiting control unit notification of sequential intent. The result is an enhanced RAID-5 design that does not suffer the write penalty and goes even further by optimizing array disk utilization for sequential workloads.

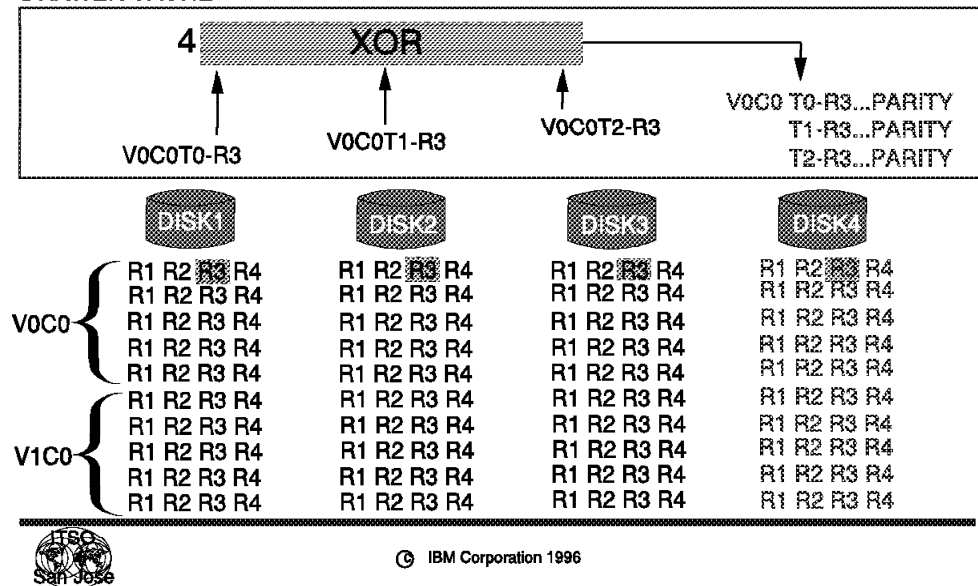
The RAMAC 2 and RAMAC 3 Array drawers use specialized microcode components to exploit these methods for update write accesses. Sometimes update write I/O is written out sequentially to the device (for example, DB2 logging). When that happens, groups of updated records are written out in sequence, and the control unit signals the sequential intent even for update write I/O. The RAMAC 2 and RAMAC 3 drawers can detect this variation of sequential access and execute a series of "SCSI block" destage maneuvers that optimize drawer operations even further.

The foil shows three records being updated sequentially in the RAMAC drawer cache—the records are logical volume 0, cylinder 0, track 0, record 3 on disk 1; logical volume 0, cylinder 0, track 1, record 3 on disk 2; and logical volume 0, cylinder 0, track 2, record 3 on disk 3. All three records are in the same slice and their ultimate locations on the three disks are highlighted. This "record arrival" pattern is similar to the previous foils on sequential I/O.

When the updated records are written to drawer cache, the traditional RAID-5 architecture would result in the staging of the old data and the old parity for each record. When the new parity was produced it would then be written out in

addition to the update record—again, this involves the write penalty of two reads followed by two writes. The next foil highlights how the RAMAC drawers react to this update write, sequential pattern to avoid the write penalty for update write I/O.

DATE	DESCRIPTION	AMOUNT	BALANCE
1980	NEW YORK	100.00	100.00
1981	NEW YORK	100.00	200.00
1982	NEW YORK	100.00	300.00
1983	NEW YORK	100.00	400.00
1984	NEW YORK	100.00	500.00
1985	NEW YORK	100.00	600.00
1986	NEW YORK	100.00	700.00
1987	NEW YORK	100.00	800.00
1988	NEW YORK	100.00	900.00
1989	NEW YORK	100.00	1000.00

DRAWER CACHE

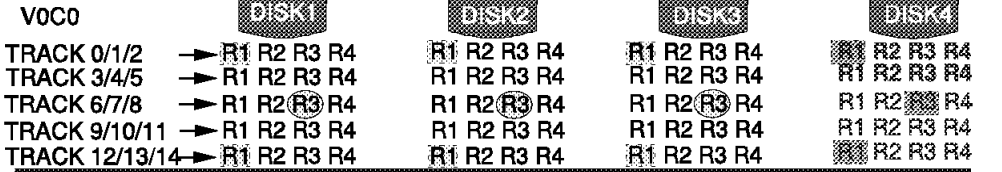
The RAMAC, RAMAC 2, and RAMAC 3 drawers have a sophisticated "operating system" that enables highly parallel, enhanced RAID-5 operations. One such component of this operating system is the background manager (BMG). The BMG scans the drawer cache contents looking for data that is in the same slice. When it detects such a pattern of updates, it initiates a destage operation based on SCSI block slices rather than track slices.

The RAMAC drawers have once again avoided the write penalty associated with RAID-5 architecture and this time for update writes as opposed to sequential writes.

DATE	DESCRIPTION	AMOUNT	CHECK	BALANCE
1999	12/31			100.00
2000	1/1			100.00
2000	1/15	50.00		50.00
2000	2/1			50.00
2000	2/15	25.00		25.00
2000	3/1			25.00
2000	3/15	12.50		12.50
2000	4/1			12.50
2000	4/15	6.25		6.25
2000	5/1			6.25
2000	5/15	3.12		3.12
2000	6/1			3.12
2000	6/15	1.56		1.56
2000	7/1			1.56
2000	7/15	0.78		0.78
2000	8/1			0.78
2000	8/15	0.39		0.39
2000	9/1			0.39
2000	9/15	0.19		0.19
2000	10/1			0.19
2000	10/15	0.09		0.09
2000	11/1			0.09
2000	11/15	0.04		0.04
2000	12/1			0.04
2000	12/15	0.02		0.02
2000	12/31			0.02

9. VOCOT14-R1

3. V0C0T8-R3
2. V0C0T7-R3
1. V0C0T6-R3



RAMAC Advanced Destage Management...

The next step is to optimize the writing of the data to the disk drives in the RAMAC drawer similar to the previous foils on sequential I/O. As before, the RAMAC drawer collects the updates in the drawer cache according to threshold management. When the threshold is reached, the drawer sorts the records into SCSI block order to minimize arm positioning and destages the data to each disk drive in a group.

The foil describes nine record updates and their order of arrival in the drawer cache. Record 3 on track 6 in cylinder 0 on volume 0 (V0C0T6-R3) is written to the drawer first. V0C0T6-R3 is written to disk 1 in the RAMAC 3 Array drawer because it resides on the third CKD track stored on this disk. If the records were destaged according to "age," record 3 on tracks 6, 7, and 8 (V0C0T6-R3, V0C0T7-R3, and V0C0T8-R3) would be destaged first followed by record 1 on tracks 0, 1, and 2 (V0C0T0-R1, V0C0T1-R1, and V0C0T2-R1), followed by record 1 on tracks 12, 13, and 14 (V0C0T12-R1, V0C0T13-R1, and V0C0T14-R1). The disk locations of these records are shaded on the foil.

The destage sequence described above would result in disk drive “arm thrashing” as the writing of the data to disk is governed by record update arrival patterns. Arm thrashing results in poor disk drive utilization and consequently less efficient destage performance.

The RAMAC 3 Array drawer avoids the performance penalty associated with disk drive arm thrashing by destaging groups of updates in “sweeps” according to SCSI block addresses. The foil shows the order in which the updated records

are written to their respective locations on the disk drives. Three records are written to each of the four disk drives (including three parity updates). V0C0T0-R1 is written to disk 1, followed by V0C0T6-R3, followed by V0C0T12-R1. This sequence enables disk 1 to satisfy the three destage operations in a single sweep in one direction. The other disks in the array execute their multiple-record destages in the same manner.

Thus, the RAMAC 3 Array drawer has optimized all aspects of high-performance RAID-5 operations—from upper interface data transfer, through hardware-assisted parity processing and logical volume mapping, to advanced destage management and intelligent disk drive utilization.

Advanced Destage Management - Summary



- Write penalty avoidance
 - Sequential processing
 - Update writes
- Disk drive optimization
 - Group destages
 - Destage on SCSI block boundaries
- Storage control-RAMAC drawer synergy
 - Exploit sequential intent indication
 - ▶ *Host software initiated*
 - ▶ *CU initiated - sequential detect*
- Exploit drawer operating system

HIGH PERFORMANCE RAID-5 OPERATIONS



© IBM Corporation 1996

Advanced Destage Management - Summary

The foil summarizes the key advantages of the RAMAC 3 drawer's advanced destage management techniques.

Write penalty avoidance

RAID-5 architecture delivers high availability through cost-effective redundancy. Multiple disks in an array formation contain data and parity information that enables data regeneration techniques to protect against data loss. However, writing data to a traditional RAID-5 disk array involves multiple I/Os (two reads and two writes) to compute the parity information that must be written with the data—therefore, the performance of traditional RAID-5 storage subsystems is significantly impacted.

The RAMAC Array family of products addresses the challenge of traditional RAID-5 performance by using a multilevel cache concept, a multiengine hardware design, a highly parallel operating system design, innovative disk drive technology and architecture, high-speed interface electronics, specialized hardware for offloading RAID-5 operations, intelligent device communications between storage control and drawer, and advanced destage management techniques. The result is an enhanced RAID-5 implementation that obliterates the write penalty for sequential workloads and certain update write I/Os. The RAMAC Array drawer's destage management implementation plays a major role in "RAID-5 write penalty avoidance."

Disk drive optimization

The improvements to the RAID-5 operations of a disk array involve avoiding reading and writing the “old data and old parity.” As mentioned above, the RAMAC 3 Array drawer implements many specialized functions that optimize RAID-5 execution. The advanced destage management functions of the RAMAC 3 Array drawer result in the most efficient use of the Ultrastar 2XP disk drives. The drawer scans for updated records in the same parity domain and schedules their destage on the basis of drawer cache thresholds and target disk locations. Updated tracks are grouped together, sorted, and destaged “back-to-back.” Whole track destages are avoided when parity can be calculated on SCSI block boundaries and the unit of destage confined to individual data sectors. The destage intelligence of the RAMAC 3 drawer results in effective and efficient drawer cache and disk drive utilization.

Storage control-RAMAC drawer synergy

The RAMAC 3 Array drawer’s destage management techniques depend on the 3990-6 and 9390 storage controls communicating the sequential nature of the workload. When the indication of sequential access is received by the drawer, it manages cache occupancy, track thresholds, and background monitoring of record updates in the same track and SCSI block slice. The drawer management tasks can be initiated by host-specified sequential intent or storage control sequential detect logic.

Exploit drawer operating system

The scheduling and initiation of destage activity from drawer cache to the Ultrastar 2XP disk drives occur at the same time as other drawer processing tasks such as data prefetch and transfer from drawer cache to storage control cache. The RAMAC 3 drawer operating system executes on multiple microprocessors that overlap many data management and movement operations. This degree of parallelism is made possible through exploitation of the multiengine hardware and highly parallel RAMAC 3 Array drawer operating system—resulting in high-performance RAID-5 operations.

Appendix B. Special Notices

This publication is intended to help customer management and technical support staff evaluate and plan for the implementation of RAMAC 3 Array Storage. The information in this publication is not intended as the specification of any programming interfaces that are provided by the 3990-6 or RAMAC Array Family of products. See the PUBLICATIONS section of the IBM Programming Announcement for 3990-6 and RAMAC 3 Array Storage for more information about what publications are considered to be product documentation.

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to the IBM Director of Licensing, IBM Corporation, 500 Columbus Avenue, Thornwood, NY 10594 USA.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any performance data contained in this document was determined in a controlled environment, and therefore, the results that may be obtained in other operating environments may vary significantly. Users of this document should verify the applicable data for their specific environment.

Reference to PTF numbers that have not been released through the normal distribution process does not imply general availability. The purpose of including these reference numbers is to alert IBM customers to specific information relative to the implementation of the PTF when it becomes available to each customer according to the normal IBM PTF distribution process.

The following terms are trademarks of the International Business Machines Corporation in the United States and/or other countries:

DB2	DFSMS
DFSMS/MVS	DFSMS/VM
DFSMSdfp	DFSMSdss
DFSMShsm	DFSMSrmm
DFSORT	ECKD
ES/3090	ES/4381
ES/9000	ESA/390
ESCON XDF	ESCON
Hardware Configuration Definition	IBM
IMS	IMS/ESA
MVS (logo)	MVS/DFP
MVS/ESA	MVS/SP
MVS/XA	NetView
No-ID Sector Format	Predictive Failure Analysis
RAMAC	S/390
S/390 Parallel Enterprise Server	Service Director
Sysplex Timer	System/390
S/390 Parallel Enterprise Server	VM/ESA
VM/XA	VSE/ESA
3090	

The following terms are trademarks of other companies:

C-bus is a trademark of Corollary, Inc.

PC Direct is a trademark of Ziff Communications Company and is used by IBM Corporation under license.

UNIX is a registered trademark in the United States and other countries licensed exclusively through X/Open Company Limited.

Microsoft, Windows, and the Windows 95 logo are trademarks or registered trademarks of Microsoft Corporation.

Other trademarks are trademarks of their respective companies.

Appendix C. Related Publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

International Technical Support Organization Publications

For information on ordering these ITSO publications see "How To Get ITSO Redbooks" on page 173.

- *IBM RAMAC Array Family*, GG24-2509
- *IBM RAMAC Array Family Additions (RAMAC 2)*, SG24-4563
- *Planning for IBM Remote Copy*, SG24-2595
- *P/DAS and Enhancements to the IBM 3990-6 and RAMAC Array Family*, SG24-4724
- *IBM 3990 Storage Control ESCON Features Presentation Guide*, GG24-3803

Redbooks on CD-ROMs

Redbooks are also available on CD-ROMs. **Order a subscription** and receive updates 2-4 times a year at significant savings.

CD-ROM Title	Subscription Number	Collection Kit Number
System/390 Redbooks Collection	SBOF-7201	SK2T-2177
Networking and Systems Management Redbooks Collection	SBOF-7370	SK2T-6022
Transaction Processing and Data Management Redbook	SBOF-7240	SK2T-8038
AS/400 Redbooks Collection	SBOF-7270	SK2T-2849
RISC System/6000 Redbooks Collection (HTML, BkMgr)	SBOF-7230	SK2T-8040
RISC System/6000 Redbooks Collection (PostScript)	SBOF-7205	SK2T-8041
Application Development Redbooks Collection	SBOF-7290	SK2T-8037
Personal Systems Redbooks Collection	SBOF-7250	SK2T-8042

Other Publications

These publications are also relevant as further information sources:

- *Remote Copy Administrator's Guide and Reference*, SC35-0169
- *ICKDSF R16 User's Guide and Reference*, GC35-0033
- *IBM 3990 Storage Control Introduction*, GA32-0098
- *IBM 3990 Operations and Recovery Guide*, GA32-0253
- *IBM 3990 Storage Control Planning, Installation, and Storage Administration Guide*, GA32-0100
- *IBM RAMAC Array Subsystem Introduction*, GC26-7004
- *IBM RAMAC Array Subsystem Reference*, GC26-7006
- *Using the IBM RAMAC Array Subsystem in an MVS, VM, or VSE Environment*, GC26-7005
- *IBM RAMAC Array DASD Introduction*, GC26-7012

- *Using the IBM RAMAC Array DASD in an MVS, VM, or VSE Environment,*
GC26-7013

How To Get ITSO Redbooks

This section explains how both customers and IBM employees can find out about ITSO redbooks, CD-ROMs, workshops, and residencies. A form for ordering books and CD-ROMs is also provided.

This information was current at the time of publication, but is continually subject to change. The latest information may be found at URL <http://www.redbooks.ibm.com>.

How IBM Employees Can Get ITSO Redbooks

Employees may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **PUBORDER** — to order hardcopies in United States
- **GOPHER link to the Internet** - type GOPHER.WTSCPOK.ITSO.IBM.COM
- **Tools disks**

To get LIST3820s of redbooks, type one of the following commands:

```
TOOLS SENDTO EHONE4 TOOLS2 REDPRINT GET SG24xxxx PACKAGE
TOOLS SENDTO CANVM2 TOOLS REDPRINT GET SG24xxxx PACKAGE (Canadian users only)
```

To get lists of redbooks:

```
TOOLS SENDTO WTSCPOK TOOLS REDBOOKS GET REDBOOKS CATALOG
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET ITSOCAT TXT
TOOLS SENDTO USDIST MKTTOOLS MKTTOOLS GET LISTSERV PACKAGE
```

To register for information on workshops, residencies, and redbooks:

```
TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ITSOREGI 1996
```

For a list of product area specialists in the ITSO:

```
TOOLS SENDTO WTSCPOK TOOLS ZDISK GET ORGCARD PACKAGE
```

- **Redbooks Home Page on the World Wide Web**
<http://w3.itso.ibm.com/redbooks>
- **IBM Direct Publications Catalog on the World Wide Web**
<http://www.elink.ibm.link.ibm.com/pb1/pb1>

IBM employees may obtain LIST3820s of redbooks from this page.

- **REDBOOKS category on INEWS**
- **Online** — send orders to: USIB6FPL at IBMMAIL or DKIBMBSH at IBMMAIL
- **Internet Listserver**

With an Internet E-mail address, anyone can subscribe to an IBM Announcement Listserver. To initiate the service, send an E-mail note to announce@webster.ibm.link.ibm.com with the keyword subscribe in the body of the note (leave the subject line blank). A category form and detailed instructions will be sent to you.

How Customers Can Get ITSO Redbooks

Customers may request ITSO deliverables (redbooks, BookManager BOOKs, and CD-ROMs) and information about redbooks, workshops, and residencies in the following ways:

- **Online Orders** (Do not send credit card information over the Internet) — send orders to:

	IBMMAIL	Internet
In United States:	usib6fpl at ibmmail	usib6fpl@ibmmail.com
In Canada:	caibmbkz at ibmmail	lmannix@vnet.ibm.com
Outside North America:	dkibmbsh at ibmmail	bookshop@dk.ibm.com

- **Telephone orders**

United States (toll free)	1-800-879-2755
Canada (toll free)	1-800-IBM-4YOU
Outside North America	(long distance charges apply)
(+45) 4810-1320 - Danish	(+45) 4810-1020 - German
(+45) 4810-1420 - Dutch	(+45) 4810-1620 - Italian
(+45) 4810-1540 - English	(+45) 4810-1270 - Norwegian
(+45) 4810-1670 - Finnish	(+45) 4810-1120 - Spanish
(+45) 4810-1220 - French	(+45) 4810-1170 - Swedish

- **Mail Orders** — send orders to:

IBM Publications Publications Customer Support P.O. Box 29570 Raleigh, NC 27626-0570 USA	IBM Publications 144-4th Avenue, S.W. Calgary, Alberta T2P 3N5 Canada	IBM Direct Services Sortemosevej 21 DK-3450 Allerød Denmark
--	--	--

- **Fax** — send orders to:

United States (toll free)	1-800-445-9269
Canada	1-403-267-4455
Outside North America	(+45) 48 14 2207 (long distance charge)

- **1-800-IBM-4FAX (United States) or (+1) 415 855 43 29 (Outside USA)** — ask for:

Index # 4421 Abstracts of new redbooks
Index # 4422 IBM redbooks
Index # 4420 Redbooks for last six months

- **Direct Services** - send note to softwareshop@vnet.ibm.com

- **On the World Wide Web**

Redbooks Home Page	http://www.redbooks.ibm.com
IBM Direct Publications Catalog	http://www.elink.ibm.link.ibm.com/pbl/pbl

- **Internet Listserver**

With an Internet E-mail address, anyone can subscribe to an IBM Announcement Listserver. To initiate the service, send an E-mail note to announce@webster.ibm.link.ibm.com with the keyword `subscribe` in the body of the note (leave the subject line blank).

IBM Redbook Order Form

Please send me the following:

Title	Order Number	Quantity

- Please put me on the mailing list for updated versions of the IBM Redbook Catalog.

First name	Last name	
Company		
Address		
City	Postal code	Country
Telephone number	Telefax number	VAT number
• Invoice to customer number _____		
• Credit card number _____		

Credit card expiration date	Card issued to	Signature
-----------------------------	----------------	-----------

We accept American Express, Diners, Eurocard, Master Card, and Visa. Payment by credit card not available in all countries. Signature mandatory for credit card payment.

DO NOT SEND CREDIT CARD INFORMATION OVER THE INTERNET.

Index

Numerics

3090

- 9390 118
 - channel 118
 - data transfer rate 118
 - ES/9000 118
 - ESCON 118
 - host support 118
 - logical paths 118
 - S/390 118

3380 track format

- 3390 track format 82, 83
- 9390-001 82
- 9390-002 83
- 9392-B33 82, 83
- cache 82, 83
- channel 82, 83
- ESCON 82, 83
- NVS 82, 83

3390 track format

- 3380 track format 82, 83
- 9390-001 82
- 9390-002 83
- 9392-B33 82, 83
- cache 82, 83
- channel 82, 83
- ESCON 82, 83
- NVS 82, 83

3390-3

- 9392-B33 24, 25
- capacity 24, 25
- disk drive 25
- logical volume 24, 25
- RAMAC 3 drawer 24, 25

3390-6

- 9390 88
- 9391-A30 88
- CE panel 88
- data transfer rate 88
- DDC cable 88
- device adapter 88
- HSDA 88
- LIC 88
- multipath 88
- PDAD 88
- RAMAC 3 drawer 88
- RAMAC 3 Storage Frame 88
- VPD 88

3990-6

- 9390 39, 40, 46, 80, 85, 86, 93, 101, 137, 140
- 9390-001 84, 99, 138
- 9390-002 99, 127
- 9391-A30 80, 101
- 9392-B33 80

3990-6 (*continued*)

- activation 99, 123
- addressing 84—86
- asynchronous destage 95
- cache 93, 95, 137, 138
- concurrent repair 101
- configuration 137, 138
- control data set 127
- data redundancy 101
- data transfer 39, 40, 46
- DB2 138
- DDC 80, 123
- delayed copyback 101
- device adapter 99
- DFP 140
- DFSMS 140
- download 99, 123
- drawer cache 46, 95
- Dynamic Sparing 101
- ESCON 106
- ESCON adapter card 127
- floating spare 101
- full frame 84, 86
- half-full frame 84, 85
- hardware assist 39, 40
- HSDA 80
- immediate copyback 101
- installation 106
- investment protection 123
- IOCP 106
- journal data set 127
- LIC 99, 123
- logical volume 80
- LRU processing 93
- migration 106
- migration consolidation 138
- MVS 140
- NVS 95, 137, 138
- NVS management 95
- parity 39, 40, 46
- partial track stage 39, 40
- performance 101, 127, 137, 138
- physical planning 106
- RAMAC 3 84—86, 95, 99, 101, 123
- RAMAC 3 Array Storage 140
- RAMAC 3 drawer 39, 40, 46, 80, 127
- RAMAC 3 Storage Frame 80, 106
- recovery site 127
- response time 138
- SDM 127
- sequential detect 46, 93
- sequential prestaging 46
- SIM 101
- software 106, 140

3990-6 (continued)

- spare drawer 101
- stage 39, 40
- state data set 127
- string 0 85, 86
- subsystem cache 95
- throughput 138
- VM 140
- VPD 93, 106
- VSE 140
- XRC 127

9390

- 3090 118
- 3390-6 88
- 3990-6 39, 40, 46, 80, 85, 86, 93, 101, 137, 140
- 9391-A30 80, 88, 101, 114
- 9392-B33 80, 114
- addressing 85, 86
- cache 93, 137
- CE panel 88
- channel 118
- concurrent repair 101
- configuration 137
- data redundancy 101
- data transfer 39, 40, 46
- data transfer rate 88, 118
- DDC 80
- DDC cable 88
- delayed copyback 101
- depth 114
- device adapter 88
- DFP 140
- DFSMS 140
- drawer cache 46
- Dynamic Sparing 101
- ES/9000 118
- ESCON 118
- floating spare 101
- full frame 86
- half-full frame 85
- hardware assist 39, 40
- height 114
- host support 118
- HSDA 80, 88
- immediate copyback 101
- LIC 88
- logical paths 118
- logical volume 80
- LRU processing 93
- multipath 88
- MVS 140
- NVS 137
- parity 39, 40, 46
- partial track stage 39, 40
- PDAD 88
- performance 101, 137
- physical specifications 114
- RAMAC 3 85, 86, 101

9390 (continued)

- RAMAC 3 Array Storage 140
- RAMAC 3 drawer 39, 40, 46, 80, 88
- RAMAC 3 Storage Frame 80, 88
- S/390 118
- sequential detect 46, 93
- sequential prestaging 46
- SIM 101
- software 140
- spare drawer 101
- stage 39, 40
- string 0 85, 86
- VM 140
- VPD 88, 93
- VSE 140
- weight 114
- width 114

9390-001

- 3380 track format 82
- 3390 track format 82
- 3990-6 84, 99, 138
- 9390-002 78, 99, 122, 126
- 9391-A30 126
- 9392-B33 82
- activation 99
- addressing 84
- basic mode 76
- cache 76, 78, 82, 122, 138
- channel 76, 78, 82, 120
- cluster symmetry 120
- configuration 138
- DB2 138
- device adapter 99
- download 99
- ESCD 110
- ESCON 76, 78, 82, 120
- ESCON adapter card 120
- extended functions 78
- feature codes 122
- field upgrade 78
- full frame 84
- half-full frame 84
- IOCP 108, 110
- LIC 99
- logical paths 120
- migration consolidation 138
- NVS 76, 78, 82, 122, 138
- performance 138
- physical ports 120
- power boundaries 116
- power supplies 78
- RAMAC 3 84, 99
- response time 138
- SDM 126
- throughput 138
- XRC 126
- XRC session 126

- 9390-002
 - 3380 track format 83
 - 3390 track format 83
 - 3990-6 99, 127
 - 9390-001 78, 99, 122, 126
 - 9391-A30 126
 - 9392-B33 83
 - activation 99
 - application system 129, 130
 - asymmetric configurations 121
 - cache 78, 83, 122
 - channel 78, 83, 121
 - cluster symmetry 121
 - control data set 127
 - device adapter 99
 - download 99
 - ESCD 113
 - ESCON 78, 83, 121
 - ESCON adapter card 121, 127, 129, 130
 - extended functions 78
 - feature codes 122
 - field upgrade 78
 - IOCP 112, 113
 - journal data set 127
 - LIC 99
 - logical paths 121
 - NVS 78, 83, 122
 - performance 127
 - physical ports 121
 - power boundaries 117
 - power supplies 78
 - PPRC 129, 130
 - RAMAC 3 99
 - RAMAC 3 drawer 127, 129, 130
 - recovery site 127
 - recovery system 129, 130
 - SDM 126, 127
 - state data set 127
 - XRC 126, 127
 - XRC session 126
- 9391-A30
 - 3390-6 88
 - 3990-6 80, 101
 - 9390 80, 88, 101, 114
 - 9390-001 126
 - 9390-002 126
 - 9392-B33 80, 114
 - CE panel 88
 - concurrent repair 101
 - copyback 100
 - data redundancy 101
 - data transfer rate 88
 - DDC 80
 - DDC cable 88
 - delayed copyback 101
 - depth 114
 - device adapter 88
 - drawer maintenance 100
- 9391-A30 (*continued*)
 - dynamic disk reconstruction 100
 - Dynamic Sparing 100, 101
 - floating spare 101
 - height 114
 - HSDA 80, 88
 - immediate copyback 101
 - LIC 88
 - logical volume 80
 - multipath 88
 - PDAD 88
 - performance 101
 - physical specifications 114
 - RAMAC 3 101
 - RAMAC 3 drawer 80, 88
 - RAMAC 3 Storage Frame 80, 88
 - SDM 126
 - SIM 101
 - spare drawer 100, 101
 - VPD 88
 - weight 114
 - width 114
 - XRC 126
 - XRC session 126
- 9392-B33
 - 3380 track format 82, 83
 - 3390 track format 82, 83
 - 3390-3 25
 - 3990-6 80
 - 9390 80, 114
 - 9390-001 82
 - 9390-002 83
 - 9391-A30 80, 114
 - cache 82, 83
 - capacity 25
 - channel 82, 83
 - DDC 80
 - depth 114
 - disk drive 25
 - ESCON 82, 83
 - height 114
 - HSDA 80
 - logical volume 25, 80
 - NVS 82, 83
 - physical specifications 114
 - RAMAC 3 drawer 25, 80
 - RAMAC 3 Storage Frame 80
 - weight 114
 - width 114
- 9393-B33
 - 3390-3 24
 - capacity 24
 - logical volume 24
 - RAMAC 3 drawer 24

A

activation

- 3990-6 99, 123
- 9390-001 99
- 9390-002 99
- DDC 123
- device adapter 99
- download 99, 123
- investment protection 123
- LIC 99, 123
- RAMAC 3 99, 123

addressing

- 3990-6 84–86
- 9390 85, 86
- 9390-001 84
- full frame 84, 86
- half-full frame 84, 85
- RAMAC 3 84–86
- string 0 85, 86

alternate sector

- CRC 56
- ID field 56
- ID sector format 56
- logical sector number 56
- MR head 56
- sector status 56
- track capacity 56
- track density 56

APAR

- AOM 141
- CUIR 142
- DEVSERV 144
- DFDSS 143
- DFSMS 144
- DFSMSdss 143
- ERE 143
- HCD 141
- ICKDSF 143
- LISTDATA 142

application system

- 9390-002 129, 130
- ESCON adapter card 129, 130
- PPRC 129, 130
- RAMAC 3 drawer 129, 130
- recovery system 129, 130

areal density

- capacity 72
- data speed 72
- defect management 72
- disk drive cache 72
- embedded sector servo 72
- MR head 72
- No-ID sector servo 72
- operating costs 72
- power management 72
- predictive failure analysis 72
- PRML channel 72
- reliability 72

areal density (*continued*)

- rotational speed 72
- track capacity 72
- track density 72
- zoned bit recording 72

array management

- background processing 35
- cache management 35
- interprocessor communications 35
- logical interface 35
- lower interface 35
- physical interface 35
- RAMAC 3 drawer 35
- resource management 35

asymmetric configurations

- 9390-002 121
- channel 121
- cluster symmetry 121
- ESCON 121
- ESCON adapter card 121
- logical paths 121
- physical ports 121

asynchronous destage

- 3990-6 95
- batch 97
- cache 95
- drawer cache 95
- large block 97
- logical volume 97
- long chain 97
- lower interface 97
- NVS 95, 97
- NVS management 95, 97
- NVS-full 97
- online 97
- performance 97
- RAMAC 3 95, 97
- RAMAC 3 drawer 97
- sequential data streams 97
- sequential detect 97
- subsystem cache 95

availability

- background task 41
- bad block 41
- capacity 48
- CRC 48
- cycle 41
- data integrity 41, 48
- destage management 48
- disk drive 48
- disk scrubbing 41
- hardware assist 48
- highly parallel 48
- logical volume 48
- mapping 48
- multiengine 48
- multitasking 48
- parity 41

availability (*continued*)

- PDAD 48
- performance 48
- RAMAC 41
- RAMAC 2 41
- RAMAC 3 41
- SCSI block 48
- sequential detect 48
- spare block 41

average seek

- data rate 68
- disk drive 68
- latency 68
- performance 68
- rpm 68
- track to track seek 68

B

background processing

- array management 35
- cache management 35
- interprocessor communications 35
- logical interface 35
- lower interface 35
- physical interface 35
- RAMAC 3 drawer 35
- resource management 35

background task

- availability 41
- bad block 41
- cycle 41
- data integrity 41
- disk scrubbing 41
- parity 41
- RAMAC 41
- RAMAC 2 41
- RAMAC 3 41
- spare block 41

bad block

- availability 41
- background task 41
- cycle 41
- data integrity 41
- disk scrubbing 41
- parity 41
- RAMAC 41
- RAMAC 2 41
- RAMAC 3 41
- spare block 41

basic mode

- 9390-001 76
- cache 76
- channel 76
- ESCON 76
- NVS 76

batch

- asynchronous destage 97
- large block 97

batch (*continued*)

- logical volume 97
- long chain 97
- lower interface 97
- NVS 97
- NVS management 97
- NVS-full 97
- online 97
- performance 97
- RAMAC 3 97
- RAMAC 3 drawer 97
- sequential data streams 97
- sequential detect 97

battery

- cache segment 66
- CRC 66
- data integrity 66
- data rate 66
- DFW 66
- disk drive 66
- disk drive cache 66
- RAMAC 3 66
- SCSI controller 66
- SCSI interface 66

C

CAA

- DASD Magic/2 132
- DCAT 132
- DSCAT 132
- modeling tools 132
- performance 132

cables

- CRC 43
- data integrity 43
- device adapter 43
- drawer 43
- frame 43
- high-speed device adapter 43
- PDAD 43
- PSM 43
- rack 43
- RAMAC 43
- RAMAC 2 43
- RAMAC 3 43

cache

- 3380 track format 82, 83
- 3390 track format 82, 83
- 3990-6 93, 95, 137, 138
- 9390 93, 137
- 9390-001 76, 78, 82, 122, 138
- 9390-002 78, 83, 122
- 9392-B33 82, 83
- asynchronous destage 95
- basic mode 76
- channel 76, 78, 82, 83
- configuration 137, 138
- DB2 138

cache (continued)

- drawer cache 95
- ESCON 76, 78, 82, 83
- extended functions 78
- feature codes 122
- field upgrade 78
- LRU processing 93
- migration consolidation 138
- NVS 76, 78, 82, 83, 95, 122, 137, 138
- NVS management 95
- performance 137, 138
- power supplies 78
- RAMAC 3 95
- RAMAC 3 drawer 22
- response time 138
- sequential detect 93
- size 8, 13, 22
- subsystem cache 95
- throughput 138
- VPD 93

cache management

- array management 35
- background processing 35
- interprocessor communications 35
- logical interface 35
- lower interface 35
- physical interface 35
- RAMAC 3 drawer 35
- resource management 35

cache manager

- drawer cache 31, 33
- driver-receiver hardware 31, 33
- HSDA 31, 33
- path manager 33
- PCAP 33
- PDAD 31, 33
- PSM 31, 33
- RAMAC 3 drawer 31, 33
- SCSI controller 33
- XOR 33

cache segment

- battery 66
- CRC 66
- data integrity 66
- data rate 66
- DFW 66
- disk drive 66
- disk drive cache 66
- RAMAC 3 66
- SCSI controller 66
- SCSI interface 66

capacity

- 3390-3 24, 25
- 9392-B33 24, 25
- areal density 72
- availability 48
- CRC 48, 64
- data integrity 48

capacity (continued)

- data speed 72
- defect management 64, 72
- destage management 48
- disk drive 25, 48
- disk drive cache 72
- ECC 64
- embedded sector servo 72
- hardware assist 48
- header data 64
- highly parallel 48
- ID field 64
- ID sector format 58
- logical volume 24, 25, 48
- mapping 48
- MR head 58, 72
- multiengine 48
- multitasking 48
- No-ID sector format 64
- No-ID sector servo 72
- operating costs 72
- PDAD 48
- performance 48, 64
- power management 64, 72
- predictive failure analysis 72
- PRML channel 72
- RAMAC 15
- RAMAC 2 15
- RAMAC 3 15
- RAMAC 3 drawer 24, 25
- RAMAC Electronic Array 15
- RAMAC Scalable Array 15
- RAMAC Virtual Array 15
- read head 58
- reliability 64, 72
- rotary actuator 58
- rotational speed 72
- SCSI block 48
- sector search 64
- sequential detect 48
- track alignment 58
- track capacity 72
- track density 72
- write head 58
- zoned bit recording 72

CE panel

- 3390-6 88
- 9390 88
- 9391-A30 88
- data transfer rate 88
- DDC cable 88
- device adapter 88
- HSDA 88
- LIC 88
- multipath 88
- PDAD 88
- RAMAC 3 drawer 88
- RAMAC 3 Storage Frame 88

CE panel *(continued)*

VPD 88

channel

3090 118

3380 track format 82, 83

3390 track format 82, 83

9390 118

9390-001 76, 78, 82, 120

9390-002 78, 83, 121

9392-B33 82, 83

asymmetric configurations 121

basic mode 76

cache 76, 78, 82, 83

cluster symmetry 120, 121

data transfer rate 118

ES/9000 118

ESCON 8, 76, 78, 82, 83, 118, 120, 121

ESCON adapter card 120, 121

extended functions 78

field upgrade 78

host support 118

logical paths 118, 120, 121

NVS 76, 78, 82, 83

physical ports 120, 121

power supplies 78

S/390 118

channel calibration

drive failures 70

drive parameters 70

idle-time 70

predictive failure analysis 70

read/write circuits 70

CKD record

destage 44

partial track 44

SCSI block 44

SCSI slice 44

sequential processing 44

update write 44

write penalty 44

CKD track

hardware assist 38

mapping 38

RAMAC 3 drawer 38

SCSI block 38

stage 38

cluster symmetry

9390-001 120

9390-002 121

asymmetric configurations 121

channel 120, 121

ESCON 120, 121

ESCON adapter card 120, 121

logical paths 120, 121

physical ports 120, 121

CMOS

data integrity 42

microcode 42

CMOS *(continued)*

performance 42

RAMAC 3 42

SCSI header 42

concurrent copy

destage 90

drawer cache 90

FPGA 90

PPRC 90

stage 90

storage frame maintenance 90

subsystem cache 90

XRC 90

concurrent repair

3990-6 101

9390 101

9391-A30 101

data redundancy 101

delayed copyback 101

drawer maintenance 103

dynamic disk reconstruction 103

Dynamic Sparing 101, 103

floating spare 101

immediate copyback 101

performance 101

RAID-5 103

RAMAC 3 101, 103

SIM 101

spare drawer 101, 103

configuration

3990-6 137, 138

9390 137

9390-001 138

cache 137, 138

DB2 138

migration consolidation 138

NVS 137, 138

performance 137, 138

response time 138

throughput 138

control data set

3990-6 127

9390-002 127

ESCON adapter card 127

journal data set 127

performance 127

RAMAC 3 drawer 127

recovery site 127

SDM 127

state data set 127

XRC 127

copyback

9390-A30 100

drawer maintenance 100

dynamic disk reconstruction 100

Dynamic Sparing 100

spare drawer 100

CRC

- alternate sector 56
- availability 48
- battery 66
- cables 43
- cache segment 66
- capacity 48, 64
- data integrity 43, 48, 66
- data rate 66
- defect management 64
- destage management 48
- device adapter 43
- DFW 66
- disk drive 48, 66
- disk drive cache 66
- drawer 43
- ECC 64
- frame 43
- hardware assist 48
- header data 64
- high-speed device adapter 43
- highly parallel 48
- ID field 56, 64
- ID sector format 56
- logical sector number 56
- logical volume 48
- mapping 48
- MR head 56
- multiengine 48
- multitasking 48
- No-ID sector format 64
- PDAD 43, 48
- performance 48, 64
- power management 64
- PSM 43
- rack 43
- RAMAC 43
- RAMAC 2 43
- RAMAC 3 43, 66
- reliability 64
- SCSI block 48
- SCSI controller 66
- SCSI interface 66
- sector search 64
- sector status 56
- sequential detect 48
- track capacity 56
- track density 56

cycle

- availability 41
- background task 41
- bad block 41
- data integrity 41
- disk scrubbing 41
- parity 41
- RAMAC 41
- RAMAC 2 41
- RAMAC 3 41

cycle (*continued*)

- spare block 41

D

DASD Magic/2

- CAA 132
- DCAT 132
- DSCAT 132
- modeling tools 132
- performance 132

data control system

- defect management 62
- disk drive 62
- LBN 62
- No-ID sector format 62
- PBN 62
- read head 62
- rotary actuator 62
- servo control system 62
- solid state memory 62
- spare block 62
- write head 62

data integrity

- availability 41, 48
- background task 41
- bad block 41
- battery 66
- cables 43
- cache segment 66
- capacity 48
- CMOS 42
- CRC 43, 48, 66
- cycle 41
- data rate 66
- destage management 48
- device adapter 43
- DFW 66
- disk drive 48, 66
- disk drive cache 66
- disk scrubbing 41
- drawer 43
- frame 43
- hardware assist 48
- high-speed device adapter 43
- highly parallel 48
- logical volume 48
- mapping 48
- microcode 42
- multiengine 48
- multitasking 48
- parity 41
- PDAD 43, 48
- performance 42, 48
- PSM 43
- rack 43
- RAMAC 41, 43
- RAMAC 2 41, 43
- RAMAC 3 41—43, 66

- data integrity (*continued*)
 - RAMAC 3 drawer 22
 - SCSI block 48
 - SCSI controller 66
 - SCSI header 42
 - SCSI interface 66
 - sequential detect 48
 - spare block 41
- data per track
 - ID sector format 61
 - No-ID sector format 61
 - track density 61
 - tracks per inch 61
- data rate
 - average seek 68
 - battery 66
 - cache segment 66
 - CRC 66
 - data integrity 66
 - DFW 66
 - disk drive 66, 68
 - disk drive cache 66
 - inner zones 54
 - latency 68
 - outer zones 54
 - performance 68
 - RAMAC 54
 - RAMAC 2 54
 - RAMAC 3 54, 66
 - rpm 68
 - SCSI controller 66
 - SCSI interface 66
 - track to track seek 68
 - zoned bit recording 54
- data redundancy
 - 3990-6 101
 - 9390 101
 - 9391-A30 101
 - concurrent repair 101
 - delayed copyback 101
 - Dynamic Sparing 101
 - floating spare 101
 - immediate copyback 101
 - performance 101
 - RAMAC 3 101
 - SIM 101
 - spare drawer 101
- data speed
 - areal density 72
 - capacity 72
 - defect management 72
 - disk drive cache 72
 - embedded sector servo 72
 - MR head 72
 - No-ID sector servo 72
 - operating costs 72
 - power management 72
 - predictive failure analysis 72

- data speed (*continued*)
 - PRML channel 72
 - reliability 72
 - rotational speed 72
 - track capacity 72
 - track density 72
 - zoned bit recording 72
- data transfer
 - 3990-6 39, 40, 46
 - 9390 39, 40, 46
 - drawer cache 46
 - hardware assist 39, 40
 - parity 39, 40, 46
 - partial track stage 39, 40
 - RAMAC 3 drawer 39, 40, 46
 - sequential detect 46
 - sequential prestaging 46
 - stage 39, 40
- data transfer rate
 - 3090 118
 - 3390-6 88
 - 9390 88, 118
 - 9391-A30 88
 - CE panel 88
 - channel 118
 - DDC cable 88
 - device adapter 88
 - ES/9000 118
 - ESCON 118
 - host support 118
 - HSDA 88
 - LIC 88
 - logical paths 118
 - multipath 88
 - PDAD 88
 - RAMAC 3 drawer 88
 - RAMAC 3 Storage Frame 88
 - S/390 118
 - VPD 88
- DB2
 - 3990-6 138
 - 9390-001 138
 - cache 138
 - configuration 138
 - migration consolidation 138
 - NVS 138
 - performance 138
 - response time 138
 - throughput 138
- DCAT
 - CAA 132
 - DASD Magic/2 132
 - DSCAT 132
 - modeling tools 132
 - performance 132
- DDC
 - 3990-6 80, 123
 - 9390 80

DDC (*continued*)

- 9391-A30 80
- 9392-B33 80
- activation 123
- download 123
- HSDA 80
- investment protection 123
- LIC 123
- logical volume 80
- RAMAC 3 123
- RAMAC 3 drawer 80
- RAMAC 3 Storage Frame 80

DDC cable

- 3390-6 88
- 9390 88
- 9391-A30 88
- CE panel 88
- data transfer rate 88
- device adapter 88
- HSDA 88
- LIC 88
- multipath 88
- PDAD 88
- RAMAC 3 drawer 88
- RAMAC 3 Storage Frame 88
- VPD 88

defect management

- areal density 72
- capacity 64, 72
- CRC 64
- data control system 62
- data speed 72
- disk drive 62
- disk drive cache 72
- ECC 64
- embedded sector servo 72
- header data 64
- ID field 64
- LBN 62
- MR head 72
- No-ID sector format 62, 64
- No-ID sector servo 72
- operating costs 72
- PBN 62
- performance 64
- power management 64, 72
- predictive failure analysis 72
- PRML channel 72
- read head 62
- reliability 64, 72
- rotary actuator 62
- rotational speed 72
- sector search 64
- servo control system 62
- solid state memory 62
- spare block 62
- track capacity 72
- track density 72

defect management (*continued*)

- write head 62
- zoned bit recording 72

delayed copyback

- 3990-6 101
- 9390 101
- 9391-A30 101
- concurrent repair 101
- data redundancy 101
- Dynamic Sparing 101
- floating spare 101
- immediate copyback 101
- performance 101
- RAMAC 3 101
- SIM 101
- spare drawer 101

depth

- 9390 114
- 9391-A30 114
- 9392-B33 114
- height 114
- physical specifications 114
- weight 114
- width 114

design

- high-availability 20
- RAID-5 20

destage

- CKD record 44
- concurrent copy 90
- drawer cache 90
- FPGA 90
- partial track 44
- PPRC 90
- SCSI block 44
- SCSI slice 44
- sequential processing 44
- stage 90
- storage frame maintenance 90
- subsystem cache 90
- update write 44
- write penalty 44
- XRC 90

destage management

- availability 48
- capacity 48
- CRC 48
- data integrity 48
- disk drive 48, 166
- drawer cache 155, 157—161, 163, 164
- drawer operating system 166
- hardware assist 48
- highly parallel 48
- logical volume 48
- mapping 48
- multiengine 48
- multitasking 48
- parity 159, 160, 163, 164

- destage management (*continued*)
 - PDAD 48
 - performance 48, 166
 - RAID-5 18, 166
 - SCSI block 48, 166
 - sequential 159
 - sequential detect 48, 166
 - sequential processing 166
 - slice 158
 - update write 161, 163, 164, 166
 - write penalty 159, 163, 166
 - XOR 159, 163, 164
- device adapter
 - 3390-6 88
 - 3990-6 99
 - 9390 88
 - 9390-001 99
 - 9390-002 99
 - 9391-A30 88
 - activation 99
 - cables 43
 - CE panel 88
 - CRC 43
 - data integrity 43
 - data transfer rate 88
 - DDC cable 88
 - download 99
 - drawer 43
 - frame 43
 - high-speed device adapter 43
 - HSDA 88
 - LIC 88, 99
 - multipath 88
 - PDAD 43, 88
 - PSM 43
 - rack 43
 - RAMAC 43
 - RAMAC 2 43
 - RAMAC 3 43, 99
 - RAMAC 3 drawer 88
 - RAMAC 3 Storage Frame 88
 - VPD 88
- DEVSERV
 - 9392-3 146
 - RTYPE 146
- DFP
 - 3990-6 140
 - 9390 140
 - DFSMS 140
 - MVS 140
 - RAMAC 3 Array Storage 140
 - software 140
- DFSMS
 - 3990-6 140
 - 9390 140
 - DFP 140
 - MVS 140
 - RAMAC 3 Array Storage 140

- DFSMS (*continued*)
 - software 140
- DFW
 - battery 66
 - cache segment 66
 - CRC 66
 - data integrity 66
 - data rate 66
 - disk drive 66
 - disk drive cache 66
 - RAMAC 3 66
 - SCSI controller 66
 - SCSI interface 66
- disaster recovery
 - fault-tolerant 148
 - performance 148
 - positioning 148
 - PPRC 148
 - RAMAC 3 Array Storage 148
 - XRC 148
- disk cache
 - MR head 52
 - predictive failure analysis 52
 - rpm 52
 - sector format 52
 - tracks per inch 52
 - zoned bit recording 52
- disk drive
 - 3.5 in. 20
 - 3390-3 25
 - 9392-B33 25
 - availability 48
 - average seek 68
 - battery 66
 - cache segment 66
 - capacity 25, 48
 - CRC 48, 66
 - data control system 62
 - data integrity 48, 66
 - data rate 66, 68
 - defect management 62
 - destage management 48, 166
 - DFW 66
 - disk drive cache 66
 - drawer operating system 166
 - hardware assist 48
 - high-performance 20
 - highly parallel 48
 - latency 68
 - LBN 62
 - logical volume 25, 48
 - mapping 48
 - multiengine 48
 - multitasking 48
 - No-ID sector format 62
 - PBN 62
 - PDAD 48
 - performance 48, 68, 166

disk drive (*continued*)

- RAID-5 166
- RAMAC 3 66
- RAMAC 3 drawer 22, 25
- read head 62
- rotary actuator 62
- rpm 68
- SCSI 20
- SCSI block 48, 166
- SCSI controller 66
- SCSI interface 66
- sequential detect 48, 166
- sequential processing 166
- servo control system 62
- solid state memory 62
- spare block 62
- track to track seek 68
- Ultrastar 20
- Ultrastar 2XP 22
- update write 166
- write head 62
- write penalty 166

disk drive cache

- areal density 72
- battery 66
- cache segment 66
- capacity 72
- CRC 66
- data integrity 66
- data rate 66
- data speed 72
- defect management 72
- DFW 66
- disk drive 66
- embedded sector servo 72
- MR head 72
- No-ID sector servo 72
- operating costs 72
- power management 72
- predictive failure analysis 72
- PRML channel 72
- RAMAC 3 66
- reliability 72
- rotational speed 72
- SCSI controller 66
- SCSI interface 66
- track capacity 72
- track density 72
- zoned bit recording 72

disk scrubbing

- availability 41
- background task 41
- bad block 41
- cycle 41
- data integrity 41
- parity 41
- RAMAC 41
- RAMAC 2 41

disk scrubbing (*continued*)

- RAMAC 3 41
- spare block 41

disk sweep

- failure analysis 71
- idle-time 71
- log recording 71

download

- 3990-6 99, 123
- 9390-001 99
- 9390-002 99
- activation 99, 123
- DDC 123
- device adapter 99
- investment protection 123
- LIC 99, 123
- RAMAC 3 99, 123

drawer

- cables 43
- CRC 43
- data integrity 43
- device adapter 43
- frame 43
- high-speed device adapter 43
- PDAD 43
- PSM 43
- rack 43
- RAMAC 43
- RAMAC 2 43
- RAMAC 3 43

drawer cache

- 3990-6 46, 95
- 9390 46
- asynchronous destage 95
- cache 95
- cache manager 31, 33
- concurrent copy 90
- data transfer 46
- destage 90
- destage management 155, 157—161, 163, 164
- driver-receiver hardware 31, 33
- FPGA 90
- HSDA 31, 33
- NVS 95
- NVS management 95
- parity 46, 159, 160, 163, 164
- path manager 33
- PCAP 33
- PDAD 31, 33
- PPRC 90
- PSM 31, 33
- RAMAC 3 95
- RAMAC 3 drawer 31, 33, 46
- SCSI controller 33
- sequential 159
- sequential detect 46
- sequential prestaging 46
- slice 158

- drawer cache (*continued*)
 - stage 90
 - storage frame maintenance 90
 - subsystem cache 90, 95
 - update write 161, 163, 164
 - write penalty 159, 163
 - XOR 33, 159, 163, 164
 - XRC 90
- drawer maintenance
 - 9390-A30 100
 - concurrent repair 103
 - copyback 100
 - dynamic disk reconstruction 100, 103
 - Dynamic Sparing 100, 103
 - RAID-5 103
 - RAMAC 3 103
 - spare drawer 100, 103
- drawer operating system
 - destage management 166
 - disk drive 166
 - performance 166
 - RAID-5 166
 - SCSI block 166
 - sequential detect 166
 - sequential processing 166
 - update write 166
 - write penalty 166
- drive failures
 - channel calibration 70
 - drive parameters 70
 - idle-time 70
 - predictive failure analysis 70
 - read/write circuits 70
- drive parameters
 - channel calibration 70
 - drive failures 70
 - idle-time 70
 - predictive failure analysis 70
 - read/write circuits 70
- driver-receiver hardware
 - cache manager 31, 33
 - drawer cache 31, 33
 - HSDA 31, 33
 - path manager 33
 - PCAP 33
 - PDAD 31, 33
 - PSM 31, 33
 - RAMAC 3 drawer 31, 33
 - SCSI controller 33
 - XOR 33
- DSCAT
 - CAA 132
 - DASD Magic/2 132
 - DCAT 132
 - modeling tools 132
 - performance 132
- dynamic disk reconstruction
 - 9390-A30 100
- dynamic disk reconstruction (*continued*)
 - concurrent repair 103
 - copyback 100
 - drawer maintenance 100, 103
 - Dynamic Sparing 100, 103
 - RAID-5 103
 - RAMAC 3 103
 - spare drawer 100, 103
- Dynamic Sparing
 - 3990-6 101
 - 9390 101
 - 9390-A30 100
 - 9391-A30 101
 - concurrent repair 101, 103
 - copyback 100
 - data redundancy 101
 - delayed copyback 101
 - drawer maintenance 100, 103
 - dynamic disk reconstruction 100, 103
 - floating spare 101
 - immediate copyback 101
 - performance 101
 - RAID-5 103
 - RAMAC 3 101, 103
 - SIM 101
 - spare drawer 100, 101, 103

E

- ECC
 - capacity 64
 - CRC 64
 - defect management 64
 - header data 64
 - ID field 64
 - No-ID sector format 64
 - performance 64
 - power management 64
 - reliability 64
 - sector search 64
- embedded sector servo
 - areal density 72
 - capacity 72
 - data speed 72
 - defect management 72
 - disk drive cache 72
 - MR head 72
 - No-ID sector servo 72
 - operating costs 72
 - power management 72
 - predictive failure analysis 72
 - PRML channel 72
 - reliability 72
 - rotational speed 72
 - track capacity 72
 - track density 72
 - zoned bit recording 72
- ES/9000
 - 3090 118

ES/9000 (*continued*)

- 9390 118
- channel 118
- data transfer rate 118
- ESCON 118
- host support 118
- logical paths 118
- S/390 118

ESCD

- 9390-001 110
- 9390-002 113
- IOCP 110, 113

ESCON

- 3090 118
- 3380 track format 82, 83
- 3390 track format 82, 83
- 3990-6 106
- 9390 118
- 9390-001 76, 78, 82, 120
- 9390-002 78, 83, 121
- 9392-B33 82, 83
- asymmetric configurations 121
- basic mode 76
- cache 76, 78, 82, 83
- channel 76, 78, 82, 83, 118, 120, 121
- cluster symmetry 120, 121
- data transfer rate 118
- ES/9000 118
- ESCON adapter card 120, 121
- extended functions 78
- field upgrade 78
- host support 118
- installation 106
- IOCP 106
- logical paths 118, 120, 121
- migration 106
- NVS 76, 78, 82, 83
- physical planning 106
- physical ports 120, 121
- power supplies 78
- RAMAC 3 Storage Frame 106
- S/390 118
- software 106
- VPD 106

ESCON adapter card

- 3990-6 127
- 9390-001 120
- 9390-002 121, 127, 129, 130
- application system 129, 130
- asymmetric configurations 121
- channel 120, 121
- cluster symmetry 120, 121
- control data set 127
- ESCON 120, 121
- journal data set 127
- logical paths 120, 121
- performance 127
- physical ports 120, 121

ESCON adapter card (*continued*)

- PPRC 129, 130
- RAMAC 3 drawer 127, 129, 130
- recovery site 127
- recovery system 129, 130
- SDM 127
- state data set 127
- XRC 127

extended functions

- 9390-001 78
- 9390-002 78
- cache 78
- channel 78
- ESCON 78
- field upgrade 78
- NVS 78
- power supplies 78

F

failure analysis

- disk sweep 71
- idle-time 71
- log recording 71

fault-tolerant

- disaster recovery 148
- performance 148
- positioning 148
- PPRC 148
- RAMAC 3 Array Storage 148
- XRC 148

feature codes

- 9390-001 122
- 9390-002 122
- cache 122
- NVS 122

field upgrade

- 9390-001 78
- 9390-002 78
- cache 78
- channel 78
- ESCON 78
- extended functions 78
- NVS 78
- power supplies 78

floating spare

- 3990-6 101
- 9390 101
- 9391-A30 101
- concurrent repair 101
- data redundancy 101
- delayed copyback 101
- Dynamic Sparing 101
- immediate copyback 101
- performance 101
- RAMAC 3 101
- SIM 101
- spare drawer 101

- floor space
 - power 115
 - RAMAC 115
 - RAMAC 2 115
 - RAMAC 3 115
- FPGA
 - concurrent copy 90
 - destage 90
 - drawer cache 90
 - PPRC 90
 - stage 90
 - storage frame maintenance 90
 - subsystem cache 90
 - XRC 90
- frame
 - cables 43
 - CRC 43
 - data integrity 43
 - device adapter 43
 - drawer 43
 - high-speed device adapter 43
 - PDAD 43
 - PSM 43
 - rack 43
 - RAMAC 43
 - RAMAC 2 43
 - RAMAC 3 43
- full frame
 - 3990-6 84
 - 9390-001 84
 - addressing 84
 - half-full frame 84
 - RAMAC 3 84

G

- general availability
 - date 8
 - RAMAC 15
 - RAMAC 2 15
 - RAMAC 3 15
 - RAMAC 3 Array Storage 8
 - RAMAC Electronic Array 15
 - RAMAC Scalable Array 15
 - RAMAC Virtual Array 15

H

- half-full frame
 - 3990-6 84—86
 - 9390 85, 86
 - 9390-001 84
 - addressing 84—86
 - full frame 84
 - RAMAC 3 84—86
 - string 0 85, 86
- hardware assist
 - 3990-6 39, 40
 - 9390 39, 40

- hardware assist (*continued*)
 - availability 48
 - capacity 48
 - CKD track 38
 - CRC 48
 - data integrity 48
 - data transfer 39, 40
 - destage management 48
 - disk drive 48
 - highly parallel 48
 - logical volume 48
 - mapping 38, 48
 - multiengine 48
 - multitasking 48
 - parity 39, 40
 - partial track stage 39, 40
 - PDAD 48
 - performance 48
 - RAMAC 3 drawer 38—40
 - SCSI block 38, 48
 - sequential detect 48
 - stage 38—40
- hardware components
 - design 18
- hardware-assisted
 - RAID-5 18
 - RAMAC 3 drawer 22
- header data
 - capacity 64
 - CRC 64
 - defect management 64
 - ECC 64
 - ID field 64
 - No-ID sector format 64
 - performance 64
 - power management 64
 - reliability 64
 - sector search 64
- height
 - 9390 114
 - 9391-A30 114
 - 9392-B33 114
 - depth 114
 - physical specifications 114
 - weight 114
 - width 114
- high-speed device adapter
 - cables 43
 - CRC 43
 - data integrity 43
 - device adapter 43
 - drawer 43
 - frame 43
 - PDAD 43
 - PSM 43
 - rack 43
 - RAMAC 43
 - RAMAC 2 43

high-speed device adapter (*continued*)

RAMAC 3 43

host support

3090 118

9390 118

channel 118

data transfer rate 118

ES/9000 118

ESCON 118

logical paths 118

S/390 118

HSDA

3390-6 88

3990-6 80

9390 80, 88

9391-A30 80, 88

9392-B33 80

cache manager 31, 33

CE panel 88

data transfer rate 88

DDC 80

DDC cable 88

device adapter 88

drawer cache 31, 33

driver-receiver hardware 31, 33

LIC 88

logical volume 80

multipath 88

path manager 33

PCAP 33

PDAD 31, 33, 88

PSM 31, 33

RAMAC 3 drawer 31, 33, 80, 88

RAMAC 3 Storage Frame 80, 88

SCSI controller 33

VPD 88

XOR 33

I

ID field

alternate sector 56

capacity 64

CRC 56, 64

defect management 64

ECC 64

header data 64

ID sector format 56, 60

logical sector number 56

MR head 56, 60

No-ID sector format 64

performance 64

power management 64

read element 60

read head trajectory 60

reliability 64

sector search 64

sector status 56

track capacity 56

ID field (*continued*)

track density 56

write element 60

ID sector format

alternate sector 56

capacity 58

CRC 56

data per track 61

ID field 56, 60

logical sector number 56

MR head 56, 58, 60

No-ID sector format 61

read element 60

read head 58

read head trajectory 60

rotary actuator 58

sector status 56

track alignment 58

track capacity 56

track density 56, 61

tracks per inch 61

write element 60

write head 58

idle-time

channel calibration 70

disk sweep 71

drive failures 70

drive parameters 70

failure analysis 71

log recording 71

predictive failure analysis 70

read/write circuits 70

immediate copyback

3990-6 101

9390 101

9391-A30 101

concurrent repair 101

data redundancy 101

delayed copyback 101

Dynamic Sparing 101

floating spare 101

performance 101

RAMAC 3 101

SIM 101

spare drawer 101

inner zones

data rate 54

outer zones 54

RAMAC 54

RAMAC 2 54

RAMAC 3 54

zoned bit recording 54

installation

3990-6 106

ESCON 106

IOCP 106

migration 106

physical planning 106

- installation (*continued*)
 - RAMAC 3 Storage Frame 106
 - software 106
 - VPD 106
- interprocessor communications
 - array management 35
 - background processing 35
 - cache management 35
 - logical interface 35
 - lower interface 35
 - physical interface 35
 - RAMAC 3 drawer 35
 - resource management 35
- investment protection
 - 3990-6 123
 - activation 123
 - DDC 123
 - download 123
 - LIC 123
 - RAMAC 3 123
- IOCP
 - 3990-6 106
 - 9390-001 108, 110
 - 9390-002 112, 113
 - ESCD 110, 113
 - ESCON 106
 - installation 106
 - migration 106
 - physical planning 106
 - RAMAC 3 Storage Frame 106
 - software 106
 - VPD 106

J

- journal data set
 - 3990-6 127
 - 9390-002 127
 - control data set 127
 - ESCON adapter card 127
 - performance 127
 - RAMAC 3 drawer 127
 - recovery site 127
 - SDM 127
 - state data set 127
 - XRC 127

L

- latency
 - average seek 68
 - data rate 68
 - disk drive 68
 - performance 68
 - rpm 68
 - track to track seek 68
- LBN
 - data control system 62
 - defect management 62

- LBN (*continued*)
 - disk drive 62
 - No-ID sector format 62
 - PBN 62
 - read head 62
 - rotary actuator 62
 - servo control system 62
 - solid state memory 62
 - spare block 62
 - write head 62
- LIC
 - 3390-6 88
 - 3990-6 99, 123
 - 9390 88
 - 9390-001 99
 - 9390-002 99
 - 9391-A30 88
 - activation 99, 123
 - CE panel 88
 - change 13
 - data transfer rate 88
 - DDC 123
 - DDC cable 88
 - device adapter 88, 99
 - download 99, 123
 - HSDA 88
 - investment protection 123
 - multipath 88
 - PDAD 88
 - RAMAC 3 99, 123
 - RAMAC 3 drawer 88
 - RAMAC 3 Storage Frame 88
 - VPD 88
- log recording
 - disk sweep 71
 - failure analysis 71
 - idle-time 71
- logical cylinder
 - logical volume 30
 - RAMAC 3 drawer 30
- logical interface
 - array management 35
 - background processing 35
 - cache management 35
 - interprocessor communications 35
 - lower interface 35
 - physical interface 35
 - RAMAC 3 drawer 35
 - resource management 35
- logical paths
 - 3090 118
 - 9390 118
 - 9390-001 120
 - 9390-002 121
 - asymmetric configurations 121
 - channel 118, 120, 121
 - cluster symmetry 120, 121
 - data transfer rate 118

logical paths *(continued)*

- ES/9000 118
- ESCON 118, 120, 121
- ESCON adapter card 120, 121
- host support 118
- physical ports 120, 121
- S/390 118

logical sector number

- alternate sector 56
- CRC 56
- ID field 56
- ID sector format 56
- MR head 56
- sector status 56
- track capacity 56
- track density 56

logical volume

- 3390-3 24, 25
- 3990-6 80
- 9390 80
- 9391-A30 80
- 9392-B33 24, 25, 80
- asynchronous destage 97
- availability 48
- batch 97
- capacity 24, 25, 48
- CRC 48
- data integrity 48
- DDC 80
- destage management 48
- disk drive 25, 48
- hardware assist 48
- highly parallel 48
- HSDA 80
- large block 97
- logical cylinder 30
- long chain 97
- lower interface 97
- mapping 26, 28, 30, 48
- multiengine 48
- multitasking 48
- NVS 97
- NVS management 97
- NVS-full 97
- online 97
- PDAD 48
- performance 48, 97
- RAMAC 2 drawer 28
- RAMAC 3 97
- RAMAC 3 drawer 22, 24, 25, 30, 80, 97
- RAMAC 3 Storage Frame 80
- RAMAC drawer 26
- SCSI block 48
- sequential data streams 97
- sequential detect 48, 97

long chain

- asynchronous destage 97
- batch 97

long chain *(continued)*

- large block 97
- logical volume 97
- lower interface 97
- NVS 97
- NVS management 97
- NVS-full 97
- online 97
- performance 97
- RAMAC 3 97
- RAMAC 3 drawer 97
- sequential data streams 97
- sequential detect 97

lower interface

- array management 35
- asynchronous destage 97
- background processing 35
- batch 97
- cache management 35
- interprocessor communications 35
- large block 97
- logical interface 35
- logical volume 97
- long chain 97
- NVS 97
- NVS management 97
- NVS-full 97
- online 97
- performance 97
- physical interface 35
- RAMAC 3 97
- RAMAC 3 drawer 35, 97
- resource management 35
- sequential data streams 97
- sequential detect 97
- speed 8

LRU processing

- 3990-6 93
- 9390 93
- cache 93
- sequential detect 93
- VPD 93

M

mapping

- availability 48
- capacity 48
- CKD track 38
- CRC 48
- data integrity 48
- destage management 48
- disk drive 48
- hardware assist 38, 48
- highly parallel 48
- logical cylinder 30
- logical volume 26, 28, 30, 48
- multiengine 48
- multitasking 48

- mapping (*continued*)
 - PDAD 48
 - performance 48
 - RAMAC 2 drawer 28
 - RAMAC 3 drawer 30, 38
 - RAMAC drawer 26
 - SCSI block 38, 48
 - sequential detect 48
 - stage 38
- microcode
 - CMOS 42
 - data integrity 42
 - performance 42
 - RAMAC 3 42
 - SCSI header 42
- migration
 - 3990-6 106
 - ESCON 106
 - installation 106
 - IOCP 106
 - physical planning 106
 - RAMAC 3 Storage Frame 106
 - software 106
 - VPD 106
- migration consolidation
 - 3990-6 138
 - 9390-001 138
 - cache 138
 - configuration 138
 - DB2 138
 - NVS 138
 - performance 138
 - response time 138
 - throughput 138
- modeling tools
 - CAA 132
 - DASD Magic/2 132
 - DCAT 132
 - DSCAT 132
 - performance 132
- MR head
 - alternate sector 56
 - areal density 72
 - capacity 58, 72
 - CRC 56
 - data speed 72
 - defect management 72
 - disk cache 52
 - disk drive cache 72
 - embedded sector servo 72
 - ID field 56, 60
 - ID sector format 56, 58, 60
 - logical sector number 56
 - No-ID sector servo 72
 - operating costs 72
 - power management 72
 - predictive failure analysis 52, 72
 - PRML channel 72

- MR head (*continued*)
 - read element 60
 - read head 58
 - read head trajectory 60
 - reliability 72
 - rotary actuator 58
 - rotational speed 72
 - rpm 52
 - sector format 52
 - sector status 56
 - track alignment 58
 - track capacity 56, 72
 - track density 56, 72
 - tracks per inch 52
 - write element 60
 - write head 58
 - zoned bit recording 52, 72
- multiengine
 - availability 48
 - capacity 48
 - CRC 48
 - data integrity 48
 - destage management 48
 - disk drive 48
 - hardware assist 48
 - highly parallel 48
 - logical volume 48
 - mapping 48
 - multitasking 48
 - PDAD 48
 - performance 48
 - SCSI block 48
 - sequential detect 48
- multipath
 - 3390-6 88
 - 9390 88
 - 9391-A30 88
 - CE panel 88
 - data transfer rate 88
 - DDC cable 88
 - device adapter 88
 - HSDA 88
 - LIC 88
 - PDAD 88
 - RAMAC 3 drawer 88
 - RAMAC 3 Storage Frame 88
 - VPD 88
- multitasking
 - availability 48
 - capacity 48
 - CRC 48
 - data integrity 48
 - destage management 48
 - disk drive 48
 - hardware assist 48
 - highly parallel 48
 - logical volume 48
 - mapping 48

multitasking (*continued*)

- multiengine 48
- PDAD 48
- performance 48
- SCSI block 48
- sequential detect 48

MVS

- 3990-6 140
- 9390 140
- DFP 140
- DFSMS 140
- RAMAC 3 Array Storage 140
- software 140

N

No-ID sector format

- capacity 64
- CRC 64
- data control system 62
- data per track 61
- defect management 62, 64
- disk drive 62
- ECC 64
- header data 64
- ID field 64
- ID sector format 61
- LBN 62
- PBN 62
- performance 64
- power management 64
- read head 62
- reliability 64
- rotary actuator 62
- sector search 64
- servo control system 62
- solid state memory 62
- spare block 62
- track density 61
- tracks per inch 61
- write head 62

No-ID sector servo

- areal density 72
- capacity 72
- data speed 72
- defect management 72
- disk drive cache 72
- embedded sector servo 72
- MR head 72
- operating costs 72
- power management 72
- predictive failure analysis 72
- PRML channel 72
- reliability 72
- rotational speed 72
- track capacity 72
- track density 72
- zoned bit recording 72

NVS

- 3380 track format 82, 83
- 3390 track format 82, 83
- 3990-6 95, 137, 138
- 9390 137
- 9390-001 76, 78, 82, 122, 138
- 9390-002 78, 83, 122
- 9392-B33 82, 83
- asynchronous destage 95, 97
- basic mode 76
- batch 97
- cache 76, 78, 82, 83, 95, 122, 137, 138
- channel 76, 78, 82, 83
- configuration 137, 138
- DB2 138
- drawer cache 95
- ESCON 76, 78, 82, 83
- extended functions 78
- feature codes 122
- field upgrade 78
- large block 97
- logical volume 97
- long chain 97
- lower interface 97
- migration consolidation 138
- NVS management 95, 97
- NVS-full 97
- online 97
- performance 97, 137, 138
- power supplies 78
- RAMAC 3 95, 97
- RAMAC 3 drawer 97
- response time 138
- sequential data streams 97
- sequential detect 97
- size 8, 13
- subsystem cache 95
- throughput 138

NVS management

- 3990-6 95
- asynchronous destage 95, 97
- batch 97
- cache 95
- drawer cache 95
- large block 97
- logical volume 97
- long chain 97
- lower interface 97
- NVS 95, 97
- NVS-full 97
- online 97
- performance 97
- RAMAC 3 95, 97
- RAMAC 3 drawer 97
- sequential data streams 97
- sequential detect 97
- subsystem cache 95

- NVS-full
 - asynchronous destage 97
 - batch 97
 - large block 97
 - logical volume 97
 - long chain 97
 - lower interface 97
 - NVS 97
 - NVS management 97
 - online 97
 - performance 97
 - RAMAC 3 97
 - RAMAC 3 drawer 97
 - sequential data streams 97
 - sequential detect 97

O

- operating costs
 - areal density 72
 - capacity 72
 - data speed 72
 - defect management 72
 - disk drive cache 72
 - embedded sector servo 72
 - MR head 72
 - No-ID sector servo 72
 - power management 72
 - predictive failure analysis 72
 - PRML channel 72
 - reliability 72
 - rotational speed 72
 - track capacity 72
 - track density 72
 - zoned bit recording 72
- outer zones
 - data rate 54
 - inner zones 54
 - RAMAC 54
 - RAMAC 2 54
 - RAMAC 3 54
 - zoned bit recording 54

P

- parity
 - 3990-6 39, 40, 46
 - 9390 39, 40, 46
 - availability 41
 - background task 41
 - bad block 41
 - cycle 41
 - data integrity 41
 - data transfer 39, 40, 46
 - destage management 159, 160, 163, 164
 - disk scrubbing 41
 - drawer cache 46, 159, 160, 163, 164
 - hardware assist 39, 40
 - parity 46, 164

- parity (*continued*)
 - partial track stage 39, 40
 - RAMAC 41
 - RAMAC 2 41
 - RAMAC 3 41
 - RAMAC 3 drawer 39, 40, 46
 - sequential 159
 - sequential detect 46
 - sequential prestaging 46
 - spare block 41
 - stage 39, 40
 - update write 163, 164
 - write penalty 159, 163
 - XOR 159, 163, 164
- partial track
 - CKD record 44
 - destage 44
 - SCSI block 44
 - SCSI slice 44
 - sequential processing 44
 - update write 44
 - write penalty 44
- partial track stage
 - 3990-6 39, 40
 - 9390 39, 40
 - data transfer 39, 40
 - hardware assist 39, 40
 - parity 39, 40
 - RAMAC 3 drawer 39, 40
 - stage 39, 40
- path manager
 - cache manager 33
 - drawer cache 33
 - driver-receiver hardware 33
 - HSDA 33
 - PCAP 33
 - PDAD 33
 - PSM 33
 - RAMAC 3 drawer 33
 - SCSI controller 33
 - XOR 33
- PBN
 - data control system 62
 - defect management 62
 - disk drive 62
 - LBN 62
 - No-ID sector format 62
 - read head 62
 - rotary actuator 62
 - servo control system 62
 - solid state memory 62
 - spare block 62
 - write head 62
- PCAP
 - cache manager 33
 - drawer cache 33
 - driver-receiver hardware 33
 - HSDA 33

PCAP (*continued*)

- path manager 33
- PDAD 33
- PSM 33
- RAMAC 3 drawer 33
- SCSI controller 33
- XOR 33

PDAD

- 3390-6 88
- 9390 88
- 9391-A30 88
- availability 48
- cables 43
- cache manager 31, 33
- capacity 48
- CE panel 88
- CRC 43, 48
- data integrity 43, 48
- data transfer rate 88
- DDC cable 88
- destage management 48
- device adapter 43, 88
- disk drive 48
- drawer 43
- drawer cache 31, 33
- driver-receiver hardware 31, 33
- frame 43
- hardware assist 48
- high-speed device adapter 43
- highly parallel 48
- HSDA 31, 33, 88
- LIC 88
- logical volume 48
- mapping 48
- multiengine 48
- multipath 88
- multitasking 48
- path manager 33
- PCAP 33
- performance 48
- PSM 31, 33, 43
- rack 43
- RAMAC 43
- RAMAC 2 43
- RAMAC 3 43
- RAMAC 3 drawer 22, 31, 33, 88
- RAMAC 3 Storage Frame 88
- SCSI block 48
- SCSI controller 33
- sequential detect 48
- VPD 88
- XOR 33

performance

- 3990-6 101, 127, 137, 138
- 9390 101, 137
- 9390-001 138
- 9390-002 127
- 9391-A30 101

performance (*continued*)

- asynchronous destage 97
- availability 48
- average seek 68
- batch 97
- CAA 132
- cache 137, 138
- capacity 48, 64
- CMOS 42
- concurrent repair 101
- configuration 137, 138
- control data set 127
- CRC 48, 64
- DASD Magic/2 132
- data integrity 42, 48
- data rate 68
- data redundancy 101
- DB2 138
- DCAT 132
- defect management 64
- delayed copyback 101
- destage management 48, 166
- disaster recovery 148
- disk drive 48, 68, 166
- drawer operating system 166
- DSCAT 132
- Dynamic Sparing 101
- ECC 64
- ESCON adapter card 127
- fault-tolerant 148
- floating spare 101
- hardware assist 48
- header data 64
- highly parallel 48
- ID field 64
- immediate copyback 101
- journal data set 127
- large block 97
- latency 68
- logical volume 48, 97
- long chain 97
- lower interface 97
- mapping 48
- microcode 42
- migration consolidation 138
- modeling tools 132
- multiengine 48
- multitasking 48
- No-ID sector format 64
- NVS 97, 137, 138
- NVS management 97
- NVS-full 97
- online 97
- PDAD 48
- positioning 148
- power management 64
- PPRC 148
- QSAM 136

performance (*continued*)

- RAID-5 166
- RAMAC 2 Array DASD 134, 136
- RAMAC 3 42, 97, 101
- RAMAC 3 Array Storage 134, 136, 148
- RAMAC 3 drawer 97, 127
- recovery site 127
- reliability 64
- response time 138
- rpm 68
- SCSI block 48, 166
- SCSI header 42
- SDM 127
- sector search 64
- sequential data streams 97
- sequential detect 48, 97, 134, 166
- sequential processing 166
- sequential write 136
- SIM 101
- spare drawer 101
- state data set 127
- throughput 138
- track to track seek 68
- update write 166
- VSAM 134
- write penalty 166
- XRC 127, 148

physical interface

- array management 35
- background processing 35
- cache management 35
- interprocessor communications 35
- logical interface 35
- lower interface 35
- RAMAC 3 drawer 35
- resource management 35

physical planning

- 3990-6 106
- ESCON 106
- installation 106
- IOCP 106
- migration 106
- RAMAC 3 Storage Frame 106
- software 106
- VPD 106

physical ports

- 9390-001 120
- 9390-002 121
- asymmetric configurations 121
- channel 120, 121
- cluster symmetry 120, 121
- ESCON 120, 121
- ESCON adapter card 120, 121
- logical paths 120, 121

physical specifications

- 9390 114
- 9391-A30 114
- 9392-B33 114

physical specifications (*continued*)

- depth 114
- height 114
- weight 114
- width 114

positioning

- disaster recovery 148
- fault-tolerant 148
- performance 148
- PPRC 148
- RAMAC 2 Array Subsystem 151
- RAMAC 3 Array Storage 148, 149
- RAMAC Electronic Array 151
- RAMAC Scalable Array 149
- RAMAC Virtual Array 149
- S/390 Internal Disk 151
- XRC 148

power

- floor space 115
- RAMAC 115
- RAMAC 2 115
- RAMAC 3 115

power boundaries

- 9390-001 116
- 9390-002 117

power management

- areal density 72
- capacity 64, 72
- CRC 64
- data speed 72
- defect management 64, 72
- disk drive cache 72
- ECC 64
- embedded sector servo 72
- header data 64
- ID field 64
- MR head 72
- No-ID sector format 64
- No-ID sector servo 72
- operating costs 72
- performance 64
- predictive failure analysis 72
- PRML channel 72
- reliability 64, 72
- rotational speed 72
- sector search 64
- track capacity 72
- track density 72
- zoned bit recording 72

power supplies

- 9390-001 78
- 9390-002 78
- cache 78
- channel 78
- ESCON 78
- extended functions 78
- field upgrade 78
- NVS 78

PPRC

- 9390-002 129, 130
- application system 129, 130
- concurrent copy 90
- destage 90
- disaster recovery 148
- drawer cache 90
- ESCON adapter card 129, 130
- fault-tolerant 148
- FPGA 90
- performance 148
- positioning 148
- RAMAC 3 Array Storage 148
- RAMAC 3 drawer 129, 130
- recovery system 129, 130
- stage 90
- storage frame maintenance 90
- subsystem cache 90
- XRC 90, 148

predictive failure analysis

- areal density 72
- capacity 72
- channel calibration 70
- data speed 72
- defect management 72
- disk cache 52
- disk drive cache 72
- drive failures 70
- drive parameters 70
- embedded sector servo 72
- idle-time 70
- MR head 52, 72
- No-ID sector servo 72
- operating costs 72
- power management 72
- PRML channel 72
- read/write circuits 70
- reliability 72
- rotational speed 72
- rpm 52
- sector format 52
- track capacity 72
- track density 72
- tracks per inch 52
- zoned bit recording 52, 72

PRML channel

- areal density 72
- capacity 72
- data speed 72
- defect management 72
- disk drive cache 72
- embedded sector servo 72
- MR head 72
- No-ID sector servo 72
- operating costs 72
- power management 72
- predictive failure analysis 72
- reliability 72

PRML channel *(continued)*

- rotational speed 72
- track capacity 72
- track density 72
- zoned bit recording 72

PSM

- cables 43
- cache manager 31, 33
- CRC 43
- data integrity 43
- device adapter 43
- drawer 43
- drawer cache 31, 33
- driver-receiver hardware 31, 33
- frame 43
- high-speed device adapter 43
- HSDA 31, 33
- path manager 33
- PCAP 33
- PDAD 31, 33, 43
- rack 43
- RAMAC 43
- RAMAC 2 43
- RAMAC 3 43
- RAMAC 3 drawer 31, 33
- SCSI controller 33
- XOR 33

PTF

- AOM 141
- CUIR 142
- DEVSERV 144
- DFDSS 143
- DFSMS 144
- DFSMSdss 143
- EREP 143
- HCD 141
- ICKDSF 143
- LISTDATA 142

Q

QSAM

- performance 136
- RAMAC 2 Array DASD 136
- RAMAC 3 Array Storage 136
- sequential write 136

R

rack

- cables 43
- CRC 43
- data integrity 43
- device adapter 43
- drawer 43
- frame 43
- high-speed device adapter 43
- PDAD 43
- PSM 43

rack (*continued*)

- RAMAC 43
- RAMAC 2 43
- RAMAC 3 43

RAID-5

- concurrent repair 103
- destage management 166
- disk drive 166
- drawer maintenance 103
- drawer operating system 166
- dynamic disk reconstruction 103
- Dynamic Sparing 103
- performance 166
- RAMAC 3 103
- SCSI block 166
- sequential detect 166
- sequential processing 166
- spare drawer 103
- update write 166
- write penalty 166

RAMAC

- availability 41
- background task 41
- bad block 41
- cables 43
- CRC 43
- cycle 41
- data integrity 41, 43
- data rate 54
- device adapter 43
- disk scrubbing 41
- drawer 43
- floor space 115
- frame 43
- high-speed device adapter 43
- inner zones 54
- outer zones 54
- parity 41
- PDAD 43
- power 115
- PSM 43
- rack 43
- RAMAC 2 41, 43, 54, 115
- RAMAC 3 41, 43, 54, 115
- spare block 41
- zoned bit recording 54

RAMAC 2

- availability 41
- background task 41
- bad block 41
- cables 43
- CRC 43
- cycle 41
- data integrity 41, 43
- data rate 54
- device adapter 43
- disk scrubbing 41
- drawer 43

RAMAC 2 (*continued*)

- floor space 115
- frame 43
- high-speed device adapter 43
- inner zones 54
- outer zones 54
- parity 41
- PDAD 43
- power 115
- PSM 43
- rack 43
- RAMAC 41, 43, 54, 115
- RAMAC 3 41, 43, 54, 115
- spare block 41
- zoned bit recording 54

RAMAC 2 Array DASD

- performance 134, 136
- QSAM 136
- RAMAC 3 Array Storage 134, 136
- sequential detect 134
- sequential write 136
- VSAM 134

RAMAC 2 Array Subsystem

- positioning 151
- RAMAC Electronic Array 151
- S/390 Internal Disk 151

RAMAC 2 drawer

- logical volume 28
- mapping 28

RAMAC 3

- 3990-6 84–86, 95, 99, 101, 123
- 9390 85, 86, 101
- 9390-001 84, 99
- 9390-002 99
- 9391-A30 101
- activation 99, 123
- addressing 84–86
- asynchronous destage 95, 97
- availability 41
- background task 41
- bad block 41
- batch 97
- battery 66
- cables 43
- cache 95
- cache segment 66
- CMOS 42
- concurrent repair 101, 103
- CRC 43, 66
- cycle 41
- data integrity 41–43, 66
- data rate 54, 66
- data redundancy 101
- DDC 123
- delayed copyback 101
- device adapter 43, 99
- DFW 66
- disk drive 66

RAMAC 3 (continued)

- disk drive cache 66
- disk scrubbing 41
- download 99, 123
- drawer 43
- drawer cache 95
- drawer maintenance 103
- dynamic disk reconstruction 103
- Dynamic Sparing 101, 103
- floating spare 101
- floor space 115
- frame 43
- full frame 84, 86
- half-full frame 84, 85
- high-speed device adapter 43
- immediate copyback 101
- inner zones 54
- investment protection 123
- large block 97
- LIC 99, 123
- logical volume 97
- long chain 97
- lower interface 97
- microcode 42
- NVS 95, 97
- NVS management 95, 97
- NVS-full 97
- online 97
- outer zones 54
- parity 41
- PDAD 43
- performance 42, 97, 101
- power 115
- PSM 43
- rack 43
- RAID-5 103
- RAMAC 41, 43, 54, 115
- RAMAC 2 41, 43, 54, 115
- RAMAC 3 drawer 97
- SCSI controller 66
- SCSI header 42
- SCSI interface 66
- sequential data streams 97
- sequential detect 97
- SIM 101
- spare block 41
- spare drawer 101, 103
- string 0 85, 86
- subsystem cache 95
- zoned bit recording 54

RAMAC 3 Array Storage

- 3990-6 140
- 9390 140
- DFP 140
- DFSMS 140
- disaster recovery 148
- fault-tolerant 148
- MVS 140

RAMAC 3 Array Storage (continued)

- performance 134, 136, 148
- positioning 148, 149
- PPRC 148
- QSAM 136
- RAMAC 2 Array DASD 134, 136
- RAMAC Scalable Array 149
- RAMAC Virtual Array 149
- sequential detect 134
- sequential write 136
- software 140
- VM 140
- VSAM 134
- VSE 140
- XRC 148

RAMAC 3 drawer

- 3390-3 24, 25
- 3390-6 88
- 3990-6 39, 40, 46, 80, 127
- 9390 39, 40, 46, 80, 88
- 9390-002 127, 129, 130
- 9391-A30 80, 88
- 9392-B33 24, 25, 80
- application system 129, 130
- asynchronous destage 97
- batch 97
- cache manager 31, 33
- capacity 24, 25
- CE panel 88
- CKD track 38
- control data set 127
- data transfer 39, 40, 46
- data transfer rate 88
- DDC 80
- DDC cable 88
- device adapter 88
- disk drive 25
- drawer cache 31, 33, 46
- driver-receiver hardware 31, 33
- ESCON adapter card 127, 129, 130
- hardware assist 38—40
- HSDA 31, 33, 80, 88
- journal data set 127
- large block 97
- LIC 88
- logical cylinder 30
- logical volume 24, 25, 30, 80, 97
- long chain 97
- lower interface 97
- mapping 30, 38
- multipath 88
- NVS 97
- NVS management 97
- NVS-full 97
- online 97
- parity 39, 40, 46
- partial track stage 39, 40
- path manager 33

RAMAC 3 drawer *(continued)*

- PCAP 33
- PDAD 31, 33, 88
- performance 97, 127
- PPRC 129, 130
- PSM 31, 33
- RAMAC 3 97
- RAMAC 3 Storage Frame 80, 88
- recovery site 127
- recovery system 129, 130
- SCSI block 38
- SCSI controller 33
- SDM 127
- sequential data streams 97
- sequential detect 46, 97
- sequential prestaging 46
- stage 38—40
- state data set 127
- VPD 88
- XOR 33
- XRC 127

RAMAC 3 Storage Frame

- 3390-6 88
- 3990-6 80, 106
- 9390 80, 88
- 9391-A30 80, 88
- 9392-B33 80
- CE panel 88
- data transfer rate 88
- DDC 80
- DDC cable 88
- device adapter 88
- ESCON 106
- HSDA 80, 88
- installation 106
- IOCP 106
- LIC 88
- logical volume 80
- migration 106
- multipath 88
- PDAD 88
- physical planning 106
- RAMAC 3 drawer 80, 88
- software 106
- VPD 88, 106

RAMAC drawer

- logical volume 26
- mapping 26

RAMAC Electronic Array

- positioning 151
- RAMAC 2 Array Subsystem 151
- S/390 Internal Disk 151

RAMAC Scalable Array

- positioning 149
- RAMAC 3 Array Storage 149
- RAMAC Virtual Array 149

RAMAC Virtual Array

- positioning 149

RAMAC Virtual Array *(continued)*

- RAMAC 3 Array Storage 149
- RAMAC Scalable Array 149

read element

- ID field 60
- ID sector format 60
- MR head 60
- read head trajectory 60
- write element 60

read head

- capacity 58
- data control system 62
- defect management 62
- disk drive 62
- ID sector format 58
- LBN 62
- MR head 58
- No-ID sector format 62
- PBN 62
- rotary actuator 58, 62
- servo control system 62
- solid state memory 62
- spare block 62
- track alignment 58
- write head 58, 62

read head trajectory

- ID field 60
- ID sector format 60
- MR head 60
- read element 60
- write element 60

read/write circuits

- channel calibration 70
- drive failures 70
- drive parameters 70
- idle-time 70
- predictive failure analysis 70

recovery

- design 18

recovery site

- 3990-6 127
- 9390-002 127
- control data set 127
- ESCON adapter card 127
- journal data set 127
- performance 127
- RAMAC 3 drawer 127
- SDM 127
- state data set 127
- XRC 127

recovery system

- 9390-002 129, 130
- application system 129, 130
- ESCON adapter card 129, 130
- PPRC 129, 130
- RAMAC 3 drawer 129, 130

reliability

- areal density 72

- reliability (*continued*)
 - capacity 64, 72
 - CRC 64
 - data speed 72
 - defect management 64, 72
 - disk drive cache 72
 - ECC 64
 - embedded sector servo 72
 - header data 64
 - ID field 64
 - MR head 72
 - No-ID sector format 64
 - No-ID sector servo 72
 - operating costs 72
 - performance 64
 - power management 64, 72
 - predictive failure analysis 72
 - PRML channel 72
 - rotational speed 72
 - sector search 64
 - track capacity 72
 - track density 72
 - zoned bit recording 72
- resource management
 - array management 35
 - background processing 35
 - cache management 35
 - interprocessor communications 35
 - logical interface 35
 - lower interface 35
 - physical interface 35
 - RAMAC 3 drawer 35
- response time
 - 3990-6 138
 - 9390-001 138
 - cache 138
 - configuration 138
 - DB2 138
 - migration consolidation 138
 - NVS 138
 - performance 138
 - throughput 138
- rotary actuator
 - capacity 58
 - data control system 62
 - defect management 62
 - disk drive 62
 - ID sector format 58
 - LBN 62
 - MR head 58
 - No-ID sector format 62
 - PBN 62
 - read head 58, 62
 - servo control system 62
 - solid state memory 62
 - spare block 62
 - track alignment 58
 - write head 58, 62

- rotational speed
 - areal density 72
 - capacity 72
 - data speed 72
 - defect management 72
 - disk drive cache 72
 - embedded sector servo 72
 - MR head 72
 - No-ID sector servo 72
 - operating costs 72
 - power management 72
 - predictive failure analysis 72
 - PRML channel 72
 - reliability 72
 - track capacity 72
 - track density 72
 - zoned bit recording 72
- rpm
 - average seek 68
 - data rate 68
 - disk cache 52
 - disk drive 68
 - latency 68
 - MR head 52
 - performance 68
 - predictive failure analysis 52
 - sector format 52
 - track to track seek 68
 - tracks per inch 52
 - zoned bit recording 52

S

- S/390
 - 3090 118
 - 9390 118
 - channel 118
 - data transfer rate 118
 - ES/9000 118
 - ESCON 118
 - host support 118
 - logical paths 118
- S/390 Internal Disk
 - positioning 151
 - RAMAC 2 Array Subsystem 151
 - RAMAC Electronic Array 151
- SCSI block
 - availability 48
 - capacity 48
 - CKD record 44
 - CKD track 38
 - CRC 48
 - data integrity 48
 - destage 44
 - destage management 48, 166
 - disk drive 48, 166
 - drawer operating system 166
 - hardware assist 38, 48
 - highly parallel 48

SCSI block *(continued)*

- logical volume 48
- mapping 38, 48
- multiengine 48
- multitasking 48
- partial track 44
- PDAD 48
- performance 48, 166
- RAID-5 166
- RAMAC 3 drawer 38
- SCSI slice 44
- sequential detect 48, 166
- sequential processing 44, 166
- stage 38
- update write 44, 166
- write penalty 44, 166

SCSI controller

- battery 66
- cache manager 33
- cache segment 66
- CRC 66
- data integrity 66
- data rate 66
- DFW 66
- disk drive 66
- disk drive cache 66
- drawer cache 33
- driver-receiver hardware 33
- HSDA 33
- path manager 33
- PCAP 33
- PDAD 33
- PSM 33
- RAMAC 3 66
- RAMAC 3 drawer 33
- SCSI interface 66
- XOR 33

SCSI header

- CMOS 42
- data integrity 42
- microcode 42
- performance 42
- RAMAC 3 42

SCSI interface

- battery 66
- cache segment 66
- CRC 66
- data integrity 66
- data rate 66
- DFW 66
- disk drive 66
- disk drive cache 66
- RAMAC 3 66
- SCSI controller 66

SCSI slice

- CKD record 44
- destage 44
- partial track 44

SCSI slice *(continued)*

- SCSI block 44
- sequential processing 44
- update write 44
- write penalty 44

SDM

- 3990-6 127
- 9390-001 126
- 9390-002 126, 127
- 9391-A30 126
- control data set 127
- ESCON adapter card 127
- journal data set 127
- performance 127
- RAMAC 3 drawer 127
- recovery site 127
- state data set 127
- XRC 126, 127
- XRC session 126

sector format

- disk cache 52
- MR head 52
- predictive failure analysis 52
- rpm 52
- tracks per inch 52
- zoned bit recording 52

sector search

- capacity 64
- CRC 64
- defect management 64
- ECC 64
- header data 64
- ID field 64
- No-ID sector format 64
- performance 64
- power management 64
- reliability 64

sector status

- alternate sector 56
- CRC 56
- ID field 56
- ID sector format 56
- logical sector number 56
- MR head 56
- track capacity 56
- track density 56

sequential

- destage management 159
- drawer cache 159
- parity 159
- write penalty 159
- XOR 159

sequential data streams

- asynchronous destage 97
- batch 97
- large block 97
- logical volume 97
- long chain 97

sequential data streams *(continued)*

- lower interface 97
- NVS 97
- NVS management 97
- NVS-full 97
- online 97
- performance 97
- RAMAC 3 97
- RAMAC 3 drawer 97
- sequential detect 97

sequential detect

- 3990-6 46, 93
- 9390 46, 93
- asynchronous destage 97
- availability 48
- batch 97
- cache 93
- capacity 48
- CRC 48
- data integrity 48
- data transfer 46
- destage management 48, 166
- disk drive 48, 166
- drawer cache 46
- drawer operating system 166
- hardware assist 48
- highly parallel 48
- large block 97
- logical volume 48, 97
- long chain 97
- lower interface 97
- LRU processing 93
- mapping 48
- multiengine 48
- multitasking 48
- NVS 97
- NVS management 97
- NVS-full 97
- online 97
- parity 46
- PDAD 48
- performance 48, 97, 134, 166
- RAID-5 166
- RAMAC 2 Array DASD 134
- RAMAC 3 97
- RAMAC 3 Array Storage 134
- RAMAC 3 drawer 46, 97
- SCSI block 48, 166
- sequential data streams 97
- sequential prestaging 46
- sequential processing 166
- update write 166
- VPD 93
- VSAM 134
- write penalty 166

sequential prestaging

- 3990-6 46
- 9390 46

sequential prestaging *(continued)*

- data transfer 46
- drawer cache 46
- parity 46
- RAMAC 3 drawer 46
- sequential detect 46

sequential processing

- CKD record 44
- destage 44
- destage management 166
- disk drive 166
- drawer operating system 166
- partial track 44
- performance 166
- RAID-5 166
- RAMAC 3 drawer 22
- SCSI block 44, 166
- SCSI slice 44
- sequential detect 166
- update write 44, 166
- write penalty 44, 166

sequential write

- performance 136
- QSAM 136
- RAMAC 2 Array DASD 136
- RAMAC 3 Array Storage 136

servo control system

- data control system 62
- defect management 62
- disk drive 62
- LBN 62
- No-ID sector format 62
- PBN 62
- read head 62
- rotary actuator 62
- solid state memory 62
- spare block 62
- write head 62

SIM

- 3990-6 101
- 9390 101
- 9391-A30 101
- concurrent repair 101
- data redundancy 101
- delayed copyback 101
- Dynamic Sparing 101
- floating spare 101
- immediate copyback 101
- performance 101
- RAMAC 3 101
- spare drawer 101

slice

- destage management 158
- drawer cache 158

software

- 3990-6 106, 140
- 9390 140
- APAR 141—144

- software (*continued*)
 - DFP 140
 - DFSMS 140
 - ESCON 106
 - installation 106
 - IOCP 106
 - migration 106
 - MVS 140
 - physical planning 106
 - PTF 141–144
 - RAMAC 3 Array Storage 140
 - RAMAC 3 Storage Frame 106
 - VM 140
 - VPD 106
 - VSE 140
- software components
 - design 18
- solid state memory
 - data control system 62
 - defect management 62
 - disk drive 62
 - LBN 62
 - No-ID sector format 62
 - PBN 62
 - read head 62
 - rotary actuator 62
 - servo control system 62
 - spare block 62
 - write head 62
- spare block
 - availability 41
 - background task 41
 - bad block 41
 - cycle 41
 - data control system 62
 - data integrity 41
 - defect management 62
 - disk drive 62
 - disk scrubbing 41
 - LBN 62
 - No-ID sector format 62
 - parity 41
 - PBN 62
 - RAMAC 41
 - RAMAC 2 41
 - RAMAC 3 41
 - read head 62
 - rotary actuator 62
 - servo control system 62
 - solid state memory 62
 - write head 62
- spare drawer
 - 3990-6 101
 - 9390 101
 - 9390-A30 100
 - 9391-A30 101
 - concurrent repair 101, 103
 - copyback 100
- spare drawer (*continued*)
 - data redundancy 101
 - delayed copyback 101
 - drawer maintenance 100, 103
 - dynamic disk reconstruction 100, 103
 - Dynamic Sparing 100, 101, 103
 - floating spare 101
 - immediate copyback 101
 - performance 101
 - RAID-5 103
 - RAMAC 3 101, 103
 - SIM 101
- stage
 - 3990-6 39, 40
 - 9390 39, 40
 - CKD track 38
 - concurrent copy 90
 - data transfer 39, 40
 - destage 90
 - drawer cache 90
 - FPGA 90
 - hardware assist 38–40
 - mapping 38
 - parity 39, 40
 - partial track stage 39, 40
 - PPRC 90
 - RAMAC 3 drawer 38–40
 - SCSI block 38
 - storage frame maintenance 90
 - subsystem cache 90
 - XRC 90
- state data set
 - 3990-6 127
 - 9390-002 127
 - control data set 127
 - ESCON adapter card 127
 - journal data set 127
 - performance 127
 - RAMAC 3 drawer 127
 - recovery site 127
 - SDM 127
 - XRC 127
- storage frame maintenance
 - concurrent copy 90
 - destage 90
 - drawer cache 90
 - FPGA 90
 - PPRC 90
 - stage 90
 - subsystem cache 90
 - XRC 90
- string 0
 - 3990-6 85, 86
 - 9390 85, 86
 - addressing 85, 86
 - full frame 86
 - half-full frame 85
 - RAMAC 3 85, 86

- subsystem cache
 - 3990-6 95
 - asynchronous destage 95
 - cache 95
 - concurrent copy 90
 - destage 90
 - drawer cache 90, 95
 - FPGA 90
 - NVS 95
 - NVS management 95
 - PPRC 90
 - RAMAC 3 95
 - stage 90
 - storage frame maintenance 90
 - XRC 90

T

- technology
 - common 20
 - RAMAC 20
 - RAMAC 2 20
 - RAMAC 3 20
- throughput
 - 3990-6 138
 - 9390-001 138
 - cache 138
 - configuration 138
 - DB2 138
 - migration consolidation 138
 - NVS 138
 - performance 138
 - response time 138
- track alignment
 - capacity 58
 - ID sector format 58
 - MR head 58
 - read head 58
 - rotary actuator 58
 - write head 58
- track capacity
 - alternate sector 56
 - areal density 72
 - capacity 72
 - CRC 56
 - data speed 72
 - defect management 72
 - disk drive cache 72
 - embedded sector servo 72
 - ID field 56
 - ID sector format 56
 - logical sector number 56
 - MR head 56, 72
 - No-ID sector servo 72
 - operating costs 72
 - power management 72
 - predictive failure analysis 72
 - PRML channel 72
 - reliability 72

- track capacity (*continued*)
 - rotational speed 72
 - sector status 56
 - track density 56, 72
 - zoned bit recording 72
- track density
 - alternate sector 56
 - areal density 72
 - capacity 72
 - CRC 56
 - data per track 61
 - data speed 72
 - defect management 72
 - disk drive cache 72
 - embedded sector servo 72
 - ID field 56
 - ID sector format 56, 61
 - logical sector number 56
 - MR head 56, 72
 - No-ID sector format 61
 - No-ID sector servo 72
 - operating costs 72
 - power management 72
 - predictive failure analysis 72
 - PRML channel 72
 - reliability 72
 - rotational speed 72
 - sector status 56
 - track capacity 56, 72
 - tracks per inch 61
 - zoned bit recording 72
- track format
 - 3380 13
 - 3390 13
- track to track seek
 - average seek 68
 - data rate 68
 - disk drive 68
 - latency 68
 - performance 68
 - rpm 68
- tracks per inch
 - data per track 61
 - disk cache 52
 - ID sector format 61
 - MR head 52
 - No-ID sector format 61
 - predictive failure analysis 52
 - rpm 52
 - sector format 52
 - track density 61
 - zoned bit recording 52

U

- update write
 - CKD record 44
 - destage 44
 - destage management 161, 163, 164, 166

update write (*continued*)

- disk drive 166
- drawer cache 161, 163, 164
- drawer operating system 166
- parity 163, 164
- partial track 44
- performance 166
- RAID-5 166
- SCSI block 44, 166
- SCSI slice 44
- sequential detect 166
- sequential processing 44, 166
- write penalty 44, 163, 166
- XOR 163, 164

V

VM

- 3990-6 140
- 9390 140
- RAMAC 3 Array Storage 140
- software 140

volume emulation

- capacity 18

VPD

- 3390-6 88
- 3990-6 93, 106
- 9390 88, 93
- 9391-A30 88
- cache 93
- CE panel 88
- data transfer rate 88
- DDC cable 88
- device adapter 88
- ESCON 106
- HSDA 88
- installation 106
- IOCP 106
- LIC 88
- LRU processing 93
- migration 106
- multipath 88
- PDAD 88
- physical planning 106
- RAMAC 3 drawer 88
- RAMAC 3 Storage Frame 88, 106
- sequential detect 93
- software 106

VSAM

- performance 134
- RAMAC 2 Array DASD 134
- RAMAC 3 Array Storage 134
- sequential detect 134

VSE

- 3990-6 140
- 9390 140
- RAMAC 3 Array Storage 140
- software 140

W

weight

- 9390 114
- 9391-A30 114
- 9392-B33 114
- depth 114
- height 114
- physical specifications 114
- width 114

width

- 9390 114
- 9391-A30 114
- 9392-B33 114
- depth 114
- height 114
- physical specifications 114
- weight 114

write element

- ID field 60
- ID sector format 60
- MR head 60
- read element 60
- read head trajectory 60

write head

- capacity 58
- data control system 62
- defect management 62
- disk drive 62
- ID sector format 58
- LBN 62
- MR head 58
- No-ID sector format 62
- PBN 62
- read head 58, 62
- rotary actuator 58, 62
- servo control system 62
- solid state memory 62
- spare block 62
- track alignment 58

write penalty

- CKD record 44
- destage 44
- destage management 159, 163, 166
- disk drive 166
- drawer cache 159, 163
- drawer operating system 166
- parity 159, 163
- partial track 44
- performance 166
- RAID-5 166
- SCSI block 44, 166
- SCSI slice 44
- sequential 159
- sequential detect 166
- sequential processing 44, 166
- update write 44, 163, 166
- XOR 159, 163

X

XOR

- cache manager 33
- destage management 159, 163, 164
- drawer cache 33, 159, 163, 164
- driver-receiver hardware 33
- HSDA 33
- parity 159, 163, 164
- path manager 33
- PCAP 33
- PDAD 33
- PSM 33
- RAMAC 3 drawer 33
- SCSI controller 33
- sequential 159
- update write 163, 164
- write penalty 159, 163

XRC

- 3990-6 127
- 9390-001 126
- 9390-002 126, 127
- 9391-A30 126
- concurrent copy 90
- control data set 127
- destage 90
- disaster recovery 148
- drawer cache 90
- ESCON adapter card 127
- fault-tolerant 148
- FPGA 90
- journal data set 127
- performance 127, 148
- positioning 148
- PPRC 90, 148
- RAMAC 3 Array Storage 148
- RAMAC 3 drawer 127
- recovery site 127
- SDM 126, 127
- stage 90
- state data set 127
- storage frame maintenance 90
- subsystem cache 90
- XRC session 126

XRC session

- 9390-001 126
- 9390-002 126
- 9391-A30 126
- SDM 126
- XRC 126

Z

zoned bit recording

- areal density 72
- capacity 72
- data rate 54
- data speed 72
- defect management 72

zoned bit recording (*continued*)

- disk cache 52
- disk drive cache 72
- embedded sector servo 72
- inner zones 54
- MR head 52, 72
- No-ID sector servo 72
- operating costs 72
- outer zones 54
- power management 72
- predictive failure analysis 52, 72
- PRML channel 72
- RAMAC 54
- RAMAC 2 54
- RAMAC 3 54
- reliability 72
- rotational speed 72
- rpm 52
- sector format 52
- track capacity 72
- track density 72
- tracks per inch 52



Printed in U.S.A.

S624-4835-00

